FORECASTING PATIENT TRAJECTORIES FROM THE EHR WITH ACTION-CONDITIONED WORLD MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce EHR-WM, a predictive world model for electronic health records that represents patient histories as action-conditioned latent trajectories. Rather than modeling clinical events solely through autoregressive token prediction, EHR-WM separates patient states from observed clinical actions and learns their evolution through a state- and action-conditioned transition operator. A state-dependent linearization predictor propagates latent changes over time, supporting multi-step rollout and action-conditioned simulation. On MIMIC-IV, generated trajectories preserve task-relevant and temporal structure, and decoded free-running rollouts remain predictive of future clinical measurements against task-matched forecasting baselines. Across classification and time-to-event tasks, EHR-WM is competitive but task-dependent relative to substantially larger EHR foundation models, using a 4.39M-parameter model over a restricted high-frequency feature space. Perturbing supplied action sequences produces corresponding changes in generated trajectories, demonstrating action-sensitive learned dynamics. These results support world modeling as a complementary framework for explicit, time-indexed modeling of patient evolution.

1 INTRODUCTION

Predictive world models have emerged as a central abstraction in reinforcement learning and robotics, where an agent learns a compact, often latent, representation of environment dynamics Ha and Schmidhuber (2018); Ding et al. (2025). Formally, given a sequence of states s_t and actions a_t, a world model parameterizes a transition operator $p_\theta(z_{t+1} \mid z_t, a_t)$, where $z_t = f_\phi(s_t)$ denotes a learned latent encoding of the observed state. This formulation enables simulation of future trajectories and supports planning through learned dynamics. Architecturally, such models combine representation learning with autoregressive temporal evolution (e.g., recurrent or transformer-based transition models). A key characteristic of this framework is the explicit factorization between controllable inputs (actions) and environment-driven evolution (states), enabling structured modeling of longitudinal systems.

A similar decomposition arises naturally in electronic health records (EHR) Evans (2016), where discrete clinical events can be viewed as partial observations of an underlying physiological trajectory. At each time step, the patient state consists of heterogeneous variables (e.g., vital signs, laboratory measurements, and demographics) that evolve over time, while treatments such as medications provide external clinical inputs that may alter subsequent evolution. This correspondence motivates casting EHR data into a (s_t, a_t) framework, enabling the learning of a latent transition model $p_\theta(z_{t+1} \mid z_t, a_t)$ that represents patient evolution jointly with observed clinical actions.

In this work, we introduce a predictive world model tailored to electronic health records, formalizing patient histories as sequences of states and clinical actions in a latent dynamical system. Existing autoregressive EHR foundation models can produce time-specific representations and generate trajectories from observed prefixes. Our distinction is therefore one of modeling objective rather than capability: EHR-WM directly parameterizes an action-conditioned transition operator over a time-indexed latent patient state. Its state-dependent linearization predictor propagates local latent changes while conditioning on observed clinical actions, providing a lightweight inductive bias for incremental patient evolution. Alternative-action rollouts characterize the response of the learned observational dynamics and should not be interpreted as causal treatment-effect estimates.

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.

We evaluate EHR-WM through latent trajectory fidelity, forecasting against observed clinical measurements, downstream classification and time-to-event prediction, and action-sequence perturbations. Forecasting comparisons include regularized linear models, GRU-D, Latent ODE, and Transformer baselines. Together, these experiments test whether a compact latent dynamics model preserves useful predictive information while exposing an explicit time-resolved representation of patient evolution.

2 Related Work

World Models Predictive world models emerged from model-based reinforcement learning, which aims to learn forward models of agent-environment dynamics Ha and Schmidhuber (2018); M. Moerland et al. (2023); Luo et al. (2024); Ding et al. (2025). Early approaches based on recurrent neural networks Rumelhart et al. (1985) focused on compressing sensory streams into predictive state representations that could be rolled forward to support planning and control Guan et al. (2023); Sekar et al. (2020). Over time, this perspective converged with advances in deep generative modeling Salakhutdinov (2015); Oussidi and Elhassouny (2018), leading to architectures that learn compact spatiotemporal representations and use them as substrates for downstream decision-making. The modern formulation popularized by Ha and Schmidhuber Ha and Schmidhuber (2018) demonstrated that learned latent dynamics could support policy optimization within imagined rollouts, separating representation and dynamics learning from downstream control. More broadly, this perspective treats predictive state modeling as a mechanism for learning representations sufficient to anticipate future observations and support sequential decision-making.

Another family of predictive models, Joint-Embedding Predictive Architectures (JEPA), has evolved from static representation learning toward modeling structured dynamics directly in latent space LeCun et al. (2022); Brotee et al. (2025). Early variants such as I-JEPA Assran et al. (2023) and subsequent extensions such as V-JEPA Bardes et al. (2024); Assran et al. (2025) demonstrate that predicting latent embeddings from context can produce high-level semantic and temporal representations without explicit reconstruction. This paradigm has since been extended across a range of applications and modalities Fei et al. (2023); Larey et al. (2026); Bardes et al. (2023); Thimonier et al. (2024); Maes et al. (2026); Luo et al. (2026); Li et al. (2026). More recent work, such as Causal JEPA Nam et al. (2026), extends this framework with intervention-aware objectives that encourage representations to distinguish changes induced by interventions from naturally occurring variation, providing a basis for reasoning over action-conditioned futures.

Recent advances such as LeWorldModel (LeWM) Maes et al. (2026) further improve the stability of latent predictive world models through regularization mechanisms such as SIGReg Balestriero and LeCun (2025), enabling end-to-end learning of predictive latent dynamics without reconstructing the observation space. Our work builds directly on this predictive representation-learning framework, but specializes it to longitudinal EHR by introducing an explicit decomposition between patient state and clinical action and a state- and action-conditioned transition operator over latent increments. Thus, the contribution is not a new general-purpose world-model training framework, but a controlled latent-dynamics formulation designed for action-conditioned patient trajectories.

EHR foundation models and longitudinal representations: Foundation models for health have emerged as a rapidly growing area of research, with large-scale pretraining increasingly applied across clinical text Lee et al. (2025b); Alsentzer et al. (2019), medical imaging Ma et al. (2024); An et al. (2025a), physiological signals Lee and Akamatsu (August 2025); Pan et al. (2026); Lee et al. (2025a; 2026), genomics Larey et al. (2026), and structured data like electronic health records Wornow et al. (2023b); Gorla et al. (2025); Gorla and Puduppully (2026) to learn transferable representations for downstream clinical tasks. Within structured EHR, foundation modeling often adopts the language-modeling abstraction, representing medical codes and clinical events as tokens within longitudinal sequences Steinberg et al. (2021). Early unsupervised approaches learned fixed-dimensional patient representations that compressed heterogeneous records into compact embeddings Miotto et al. (2016). More recently, Transformer-based models have adapted BERT- and GPT-style pretraining objectives to structured EHR, producing representations that support transfer across a wide range of downstream clinical prediction tasks Devlin et al. (2019); Li et al. (2020); Rasmy et al. (2021); Odgaard et al. (2024); An et al. (2025b). Subsequent approaches incorporate temporal information, visit structure, and event timing directly into the representation while retaining a general-purpose pretrained interface for downstream prediction Pang et al. (2021).

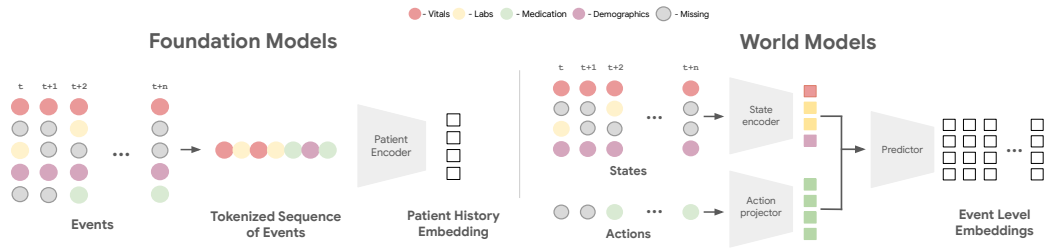

Figure 1: **From sequence-based EHR representations to trajectory-based world models.** Foundation models commonly represent electronic health records as longitudinal sequences of heterogeneous events (e.g., vitals, labs, medications) and expose learned representations of the observed patient history for downstream prediction. In contrast, our world model formulation explicitly decomposes the record into patient states and clinical actions and models their evolution within a controlled latent trajectory system. This yields time-indexed latent states whose transitions are directly conditioned on observed clinical actions, enabling trajectory forecasting and intervention-conditioned simulation.

Autoregressive EHR models use next-event or time-to-event prediction to learn transferable longitudinal representations Steinberg et al. (2021); Wornow et al. (2023a); Steinberg et al. (2024); Jing et al. (2026); Shmatko et al. (2025); Zhang and Li (2025). These models can expose representations at intermediate timepoints and generate future trajectories from observed prefixes or inserted events Shmatko et al. (2025). Our distinction is therefore one of objective: rather than modeling future clinical events or tokens, EHR-WM directly parameterizes the evolution of a latent patient state under an explicit clinical action sequence.

Concurrent healthcare world-model work has also explored long-horizon patient trajectory generation and world-model training for longitudinal EHR Mu et al. (2026); Wang et al. (2026a); Lutsker et al. (2026), though none have been formally published. Our focus is narrower: we explicitly factor observed patient state from clinical action and model action-conditioned evolution in longitudinal inpatient EHR time series.

3 METHODS

Cohort construction and temporal discretization We construct longitudinal patient trajectories from MIMIC-IV Johnson et al. (2023) using the MEDS schema Arnrich et al. (2024); McDermott et al. (2025) and partition patients into disjoint training, validation, and test cohorts. To prevent information leakage, all preprocessing statistics, including normalization parameters and feature frequencies, are computed on the training split and applied unchanged to validation and test [1].

The inpatient setting provides a useful starting point for trajectory-based world modeling because patient state is observed and acted upon at relatively high temporal frequency. Hospitalized patients undergo repeated measurement of vital signs and laboratory values, while medications and other treatments are administered throughout the admission in response to evolving clinical status. This produces comparatively dense sequences of both physiological observations and clinical actions, making inpatient EHR a useful environment for studying whether a model can learn state transitions that remain sensitive to changes in treatment. At the same time, this setting should be viewed as a controlled starting point rather than a complete representation of longitudinal healthcare, since outpatient and chronic disease trajectories are substantially sparser and unfold over much longer timescales (Wang et al., 2026a).

We impose a discrete-time structure by partitioning each patient timeline into fixed intervals of $\Delta = 6$ hours (Experiment in Appendix Section G.3). Within each interval, multiple observations of the same continuous variable are aggregated via the mean. Continuous features are standardized via z-score normalization computed on observed values. Each variable is then represented by a two-channel encoding consisting of its normalized value and a corresponding binary observation mask to encode missingness. For timesteps in which a feature is unobserved, its normalized value is set to zero

[1] https://github.com/Medical-Event-Data-Standard

and its mask to zero, ensuring that missingness is explicitly identifiable rather than conflated with mean-valued observations. This yields a regularly sampled sequence indexed by $t \in \{1, \ldots, T\}$.

State and action representation At each timestep t, we construct a state vector x_t by concatenating clinical features derived from continuous features including vital signs, laboratory measurements, and static features including demographics. These collectively represent the patient's physiological state and clinical context. Observed medication administrations within each time window are modeled explicitly as a binary clinical action vector $a_t \in \{0, 1\}^{d_a}$. The resulting sequence $\{(x_t, a_t)\}_{t=1}^{T}$ defines a discrete-time controlled trajectory.

Feature space restriction To ensure a consistent representation across a large and heterogeneous population, we restrict the feature space to the top-k most frequent variables computed on the training corpus. This selection is applied separately across continuous variables (e.g., vitals and labs) and intervention codes (e.g., medications). Limiting to a shared top-k feature set mitigates high-cardinality effects and reduces sparsity, yielding a fixed-dimensional representation that is stable across patients and scalable to large cohorts.

Sequence construction Each patient trajectory is represented as a sequence $\{(x_t, a_t)\}_{t=1}^{T}$ over discretized time, providing a structured and uniformly sampled approximation of the underlying clinical process suitable for downstream sequential modeling.

3.1 State-Dependent Linearization World Model

We adopt the LeWorldModel framework Maes et al. (2026), a latent predictive modeling paradigm inspired by JEPA-style objectives, and specialize it to the clinical setting by introducing a controlled dynamical system over patient trajectories. The central component of our formulation is a *state-dependent linearization predictor*, which models clinical actions as inputs that modulate latent transition dynamics.

Latent state construction Observations $x_t \in \mathbb{R}^{D_x}$ are mapped to a latent state $z_t \in \mathbb{R}^{D}$ via an encoder f_ϕ,

$$z_t = f_\phi(x_t), \tag{1}$$

where f_ϕ is a feedforward projection network. The learned representation z_t provides a compact predictive state from which subsequent latent evolution is modeled. We use a latent dimensionality of $D = 128$, following the scale of related latent-dynamics models Wang et al. (2026b). Appendix Section G.1 further analyzes the effective rank and spectral structure of the learned transition operators, showing that the model utilizes a substantial fraction of the available latent dimensions.

Controlled latent dynamics via state-dependent linearization. We model temporal evolution as a controlled discrete-time dynamical system over latent increments. Given a latent state $z_t \in \mathbb{R}^D$ and clinical action a_t, the predictor outputs a state- and action-dependent linear operator

$$F_\theta(z_t, a_t) \in \mathbb{R}^{D \times D}, \tag{2}$$

which parameterizes the local propagation of latent changes.

Let $\Delta z_t = z_t - z_{t-1}$. The next state is predicted as

$$\hat{z}_{t+1} = z_t + F_\theta(z_t, a_t)\, \Delta z_t, \tag{3}$$

which propagates the previous latent displacement through a state- and action-dependent linear map. Since Δz_1 depends on the unobserved z_0, we initialize the dynamics at $t = 2$ and apply the update for $t \geq 2$.

This parameterization is motivated by local linearization of smooth dynamical systems. For a latent transition map G, small perturbations around the current state can be approximated by

$$\delta z_{t+1} \approx J_z G(z_t, a_t)\, \delta z_t, \tag{4}$$

where $J_z G(z_t, a_t)$ denotes the Jacobian of the transition with respect to the latent state at a fixed action. We therefore learn $F_\theta(z_t, a_t)$ as a state- and action-dependent approximation to such a local transition operator, using the observed latent increment Δz_t as the propagated displacement. Conditioning the operator on a_t allows different clinical action sequences to induce different local transition dynamics.

The resulting update is multiplicative in Δz_t and does not include a separate additive action-forcing term. Consequently, the formulation places an inductive bias on actions modulating existing latent dynamics; the associated stationary-state limitation is discussed further in the Appendix. With respect to z_t alone, the recurrence is second-order because it depends on z_{t-1} through Δz_t. Equivalently, it can be written as a first-order controlled system in the augmented state $(z_t, \Delta z_t)$, conditioned on a_t.

Latent predictive objective Given a trajectory $(x_{t:t+T}, a_{t:t+T})$, the encoder produces latent states $z_t = f_\phi(x_t)$ and the predictor recursively generates $\hat{z}_{t+j}$ for $j \geq 1$. The model is trained using a latent prediction loss

$$\mathcal{L}_{\text{pred}} = \frac{1}{k} \sum_{j=1}^{k} \|\hat{z}_{t+j} - z_{t+j}\|_2^2 , \tag{5}$$

which enforces consistency between predicted and encoder-derived future states without requiring reconstruction of the observation space. Gradients propagate through both predicted and encoder-derived latent states seen commonly in the JEPA paradigm. This objective encourages the latent representation and transition model to support predictable multi-step evolution from the current augmented latent state under the observed sequence of clinical actions.

To regularize the geometry of the latent space, LeWorldModel adopts Sketched Isotropic Gaussian Regularization (SIGReg) Balestriero and LeCun (2025), enforcing that the empirical latent distribution matches $\mathcal{N}(0, I)$ via characteristic function matching. The full objective is

$$\mathcal{L} = \mathcal{L}_{\text{pred}} + \lambda_{\text{sig}} \mathcal{L}_{\text{SIGReg}}. \tag{6}$$

Curriculum learning and training dynamics Learning long-horizon controlled dynamics directly can be unstable due to compounding prediction error, particularly in noisy and partially observed clinical trajectories. We therefore employ a curriculum over prediction horizon and input distribution, combining teacher forcing with a sliding-window training scheme Lamb et al. (2016); Bengio et al. (2009); Zhou et al. (2026). At training time, the predictor receives either the encoder-derived latent z_t or its own prediction $\hat{z}_t$, sampled with probability ρ and $1 - \rho$, respectively. This defines a stochastic rollout state

$$\tilde{z}_t = \begin{cases} z_t & \text{with probability } \rho, \\ \hat{z}_t & \text{with probability } 1 - \rho. \end{cases} \tag{7}$$

We jointly increase the prediction horizon k while annealing ρ, transitioning from fully teacher-forced one-step prediction to partially autoregressive multi-step rollouts. Training is performed over sliding windows of trajectories, allowing the model to observe diverse local contexts while gradually extending to longer effective horizons. This curriculum first emphasizes accurate local transitions and subsequently trains stable composition of the learned transition operators over longer rollouts. All optimization settings and training hyperparameters are described in Appendix E.

4 Experiments

Trajectory fidelity and latent generative validity We first evaluate the extent to which the learned world model preserves the structure and task-relevant information of encoder-derived patient trajectories during autoregressive rollout. Here, the central question is whether predicted latent trajectories $\{\hat{z}_t\}$ generated by the world model remain aligned with the encoder-derived trajectories $\{z_t\}$ that define the model's latent prediction target. We emphasize that this analysis measures *latent trajectory fidelity*: agreement between $\hat{z}_t$ and z_t does not by itself establish fidelity to the underlying clinical observations. We therefore complement these analyses with direct decoding and forecasting in the original clinical feature space in the subsequent section.

Within the latent space, we evaluate rollout fidelity along several complementary axes, including preservation of downstream predictive signal, temporal ordering, and trajectory-level distributional structure. Rather than treating any single metric as evidence of exact patient reconstruction, these analyses test whether free-running rollouts preserve properties of the encoder-derived trajectories despite compounding prediction error.

Latent Feature Fidelity: We first consider a suite of binary classification and time-to-event (TTE) prediction tasks spanning acute conditions, chronic diseases, and utilization outcomes Wornow et al. (2023a). Task definitions, prediction horizons, and cohort construction are provided in Appendix Section D. For each task, we apply the same downstream probe independently to encoder-derived latent states z_t, denoted ORACLE, and model-generated latent states $\hat{z}_t$, denoted ROLLOUT. Probes are evaluated at each valid prediction timepoint along the patient trajectory, where valid timepoints are restricted to those preceding outcome occurrence and for which the patient remains at risk. Performance is averaged across timepoints using the same protocol for Oracle and Rollout representations. This paired evaluation isolates the loss of task-relevant information attributable to autoregressive rollout: if the learned predictor closely preserves the encoder-derived representation, predictive performance should remain similar between the two trajectories.

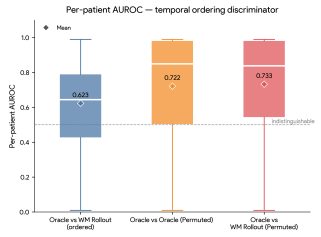

Figure 2: **Discriminability of encoder-derived and WM-generated latent trajectories.** A discriminator is trained to distinguish encoder-derived trajectories from world-model rollouts. Performance approaching chance (AUROC $\approx$ 0.5) indicates limited separability in latent space. Temporal permutation increases discriminability, providing a complementary test that the learned representations contain non-exchangeable temporal structure.

The results, summarized in Table 1, show that ROLLOUT and ORACLE representations generally retain similar downstream predictive performance across the evaluated tasks, although the magnitude and direction of the difference are task-dependent. The comparison demonstrates that multi-step rollout preserves much of the task-relevant information contained in the latent representation.For imbalanced classification outcomes, we additionally report prevalence and AUPRC in Appendix G.4; Appendix G.5 separately evaluates calibration of downstream binary predictions across rollout horizons.

Temporal Structure Preservation We next probe whether the learned world model preserves temporal structure using a discriminator-based evaluation between encoder-derived ORACLE trajectories and predicted ROLLOUT trajectories. We consider three comparisons: (i) ordered ORACLE trajectories versus ordered ROLLOUT trajectories, (ii) ordered ORACLE trajectories versus randomly permuted versions of the same trajectories, and (iii) ordered ORACLE trajectories versus randomly permuted ROLLOUT trajectories. The latter two explicitly perturb temporal order while preserving the underlying set of latent states, allowing us to test whether sequence structure itself contributes to discriminability.

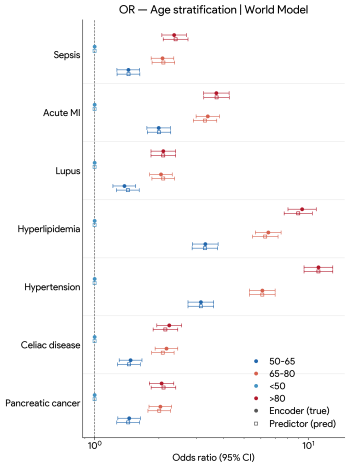

Figure 3: **Age-stratified odds ratios from generated trajectories.** Rollout-derived odds ratios (95% CI) follow similar age-stratified association patterns to encoder-derived representations.

As shown in Figure 2, discriminability is lowest when both ORACLE and ROLLOUT trajectories retain their original temporal ordering, and increases when either sequence is randomly permuted. Because the permutation operation preserves the marginal collection of latent states while disrupting their temporal arrangement, this asymmetry indicates that the learned trajectories preserve non-exchangeable temporal structure. In other words, the predictor is not merely reproducing an unordered distribution of plausible latent states; it maintains temporal dependencies that are disrupted when the sequence is shuffled. We interpret this as evidence that the rollout dynamics retain directional structure in the learned patient representation.

Subgroup Structure Preservation Lastly, we evaluate whether generated trajectories preserve clinically relevant subgroup structure by examining age-stratified odds ratios (ORs; Figure 3). Across the evaluated conditions, rollout-derived estimates broadly track those from encoder-derived representations, with similar trends across age groups. This suggests that subgroup-associated statistical structure present in the encoder representation is retained under multi-step rollout, providing comple-

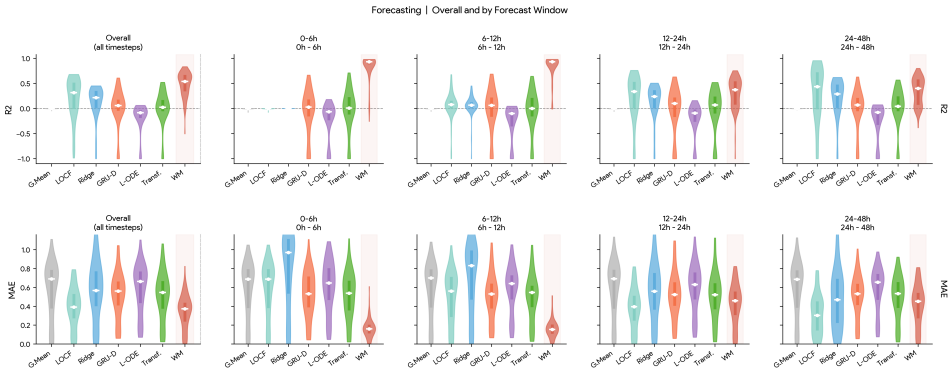

Figure 4: **Forecasting performance across prediction windows (0–6h, 6–12h, 12–24h, 24–48h) and aggregated over 80 continuous clinical variables.** We report R^2 (top row) and MAE (bottom row) for global mean, LOCF, Ridge Regression, GRU-D, Latent ODE, Transformer and EHR-WM baselines; violin plots denote the distribution across variables.

mentary evidence that the learned dynamics preserve structured clinical information beyond pointwise latent similarity. This analysis does not establish physiological or causal validity.

Forecasting via rollout decoding. To evaluate whether EHR-WM generates trajectories that remain predictive in the original clinical feature space, we adopt a rollout-based decoding protocol that directly compares model-generated forecasts with observed future measurements. Starting from the initial observed trajectory context, we encode the patient state into latent representations and iteratively apply the learned transition dynamics to obtain future latent states $\{\hat{z}_{t+\tau}\}_{\tau=1}^{H}$. At each rollout step, a shared linear decoder maps $\hat{z}_{t+\tau}$ to predictions $\hat{x}_{t+\tau}$ over the original continuous clinical variables. After initialization, the model is rolled forward without access to ground-truth future states, such that errors in the learned transition dynamics propagate through subsequent predictions.

This evaluation complements the latent-space fidelity analyses above by measuring forecasting performance directly against observed clinical values rather than encoder-derived representations. It therefore tests whether the temporal information preserved by the latent dynamics remains recoverable in the original observation space. We compare EHR-WM against task-matched forecasting baselines, including global mean, last observation carried forward (LOCF), ridge regression, GRU-D Che et al. (2016), Latent ODE Rubanova et al. (2019a), and a Transformer forecaster Vaswani et al. (2023), using the same input context, prediction horizons, target variables, and evaluation protocol.

Figure 4 summarizes performance across 80 continuous clinical variables using coefficient of determination (R^2) and mean absolute error (MAE) over increasing forecast horizons. Each violin plot shows the distribution of per-variable performance, allowing both central tendency and variability across clinical measurements to be assessed. EHR-WM remains competitive across horizons and maintains informative forecasts as rollout depth increases, although performance varies substantially across variables and prediction windows. Short-horizon prediction is comparatively favorable for persistence-based methods such as LOCF, whereas the relative behavior of the models changes as the forecast horizon increases and persistence becomes a weaker assumption.

Because evaluation is performed on free-running multi-step rollouts, errors introduced at one step affect subsequent predictions. The resulting degradation with forecast horizon therefore reflects both uncertainty in future clinical evolution and compounding transition error. Rather than interpreting performance at any single horizon as evidence of exact trajectory reconstruction, we use the combined R^2 and MAE profiles to characterize how well each model preserves clinically observed variation under increasingly long autonomous rollouts. Detailed per-variable results, including stratification by horizon, are provided in Appendix Sections G.8 and G.9.

These results provide a direct observation-space complement to the latent fidelity experiments: the learned trajectories not only preserve information within the encoder representation, but also support recovery of future clinical measurements under multi-step rollout.

Benchmark comparison across modeling paradigms We next compare EHR-WM against established baselines across binary classification and time-to-event prediction. These include traditional tabular methods (XGBoost Chen and Guestrin (2016)), large-scale EHR foundation models (CLMBR

Table 1: **Comparison across classification and survival tasks with prevalence and WM rollout/oracle decomposition.** We augment the evaluation by reporting task prevalence and decomposing EHR-WM into rollout (predicted latent trajectory) and oracle (encoder-derived latent trajectory from observed patient states) variants, isolating representation quality from rollout error.

		Binary Classification (AUC)					Time-to-Event (C-Index)				
Task	Prev	GBM	CLMBR	MOTOR	EHR WM		CoxPH	DeepHit	DeepSurv	EHR WM	
					Rollout	Oracle				Rollout	Oracle
Param Size		–	141M	143M	4.39M		–	145k	133k	4.39M	
Acute MI	4.9%	$0.723_{\pm 0.015}$	$0.726_{\pm 0.013}$	$0.739_{\pm 0.014}$	$\mathbf{0.744}_{\pm 0.012}$	$0.756_{\pm 0.015}$	$\mathbf{0.744}_{\pm 0.013}$	$0.707_{\pm 0.016}$	$0.536_{\pm 0.018}$	$0.733_{\pm 0.013}$	$0.746_{\pm 0.014}$
Lupus	0.4%	$0.711_{\pm 0.022}$	$0.741_{\pm 0.018}$	$\mathbf{0.755}_{\pm 0.017}$	$0.631_{\pm 0.019}$	$0.620_{\pm 0.020}$	$0.574_{\pm 0.021}$	$\mathbf{0.583}_{\pm 0.018}$	$0.513_{\pm 0.023}$	$\mathbf{0.630}_{\pm 0.017}$	$0.620_{\pm 0.018}$
Hyperlipidemia	30.9%	$0.666_{\pm 0.010}$	$0.671_{\pm 0.009}$	$0.673_{\pm 0.010}$	$\mathbf{0.678}_{\pm 0.008}$	$0.680_{\pm 0.009}$	$0.560_{\pm 0.011}$	$0.557_{\pm 0.010}$	$0.579_{\pm 0.009}$	$\mathbf{0.624}_{\pm 0.008}$	$0.628_{\pm 0.008}$
Hypertension	45.2%	$0.633_{\pm 0.009}$	$0.711_{\pm 0.008}$	$0.719_{\pm 0.008}$	$\mathbf{0.734}_{\pm 0.007}$	$0.740_{\pm 0.008}$	$0.562_{\pm 0.010}$	$0.541_{\pm 0.011}$	$\mathbf{0.571}_{\pm 0.008}$	$0.623_{\pm 0.007}$	$0.627_{\pm 0.007}$
Celiac	0.5%	$\mathbf{0.729}_{\pm 0.021}$	$0.554_{\pm 0.020}$	$0.631_{\pm 0.017}$	$0.627_{\pm 0.018}$	$0.621_{\pm 0.019}$	$0.527_{\pm 0.022}$	$0.514_{\pm 0.021}$	$0.543_{\pm 0.018}$	$\mathbf{0.629}_{\pm 0.017}$	$0.623_{\pm 0.017}$
Pancreatic Cancer	0.8%	$0.811_{\pm 0.019}$	$0.809_{\pm 0.018}$	$\mathbf{0.831}_{\pm 0.017}$	$0.740_{\pm 0.020}$	$0.749_{\pm 0.019}$	$\mathbf{0.863}_{\pm 0.015}$	$0.725_{\pm 0.018}$	$0.599_{\pm 0.021}$	$0.740_{\pm 0.017}$	$0.750_{\pm 0.016}$
ICU Admission	20.9%	$0.771_{\pm 0.011}$	$0.831_{\pm 0.010}$	$0.865_{\pm 0.009}$	$\mathbf{0.891}_{\pm 0.008}$	$0.894_{\pm 0.009}$	$0.748_{\pm 0.011}$	$0.650_{\pm 0.013}$	$0.594_{\pm 0.014}$	$\mathbf{0.828}_{\pm 0.009}$	$0.827_{\pm 0.009}$
Long LOS	15.0%	$0.777_{\pm 0.012}$	$0.809_{\pm 0.011}$	$\mathbf{0.821}_{\pm 0.010}$	$0.768_{\pm 0.013}$	$0.746_{\pm 0.014}$	$0.407_{\pm 0.017}$	$0.501_{\pm 0.015}$	$\mathbf{0.596}_{\pm 0.012}$	$0.385_{\pm 0.018}$	$0.389_{\pm 0.017}$
30-day Readmission	13.2%	$0.712_{\pm 0.013}$	$\mathbf{0.777}_{\pm 0.011}$	$0.738_{\pm 0.012}$	$0.622_{\pm 0.014}$	$0.627_{\pm 0.014}$	$0.721_{\pm 0.012}$	$0.806_{\pm 0.010}$	$0.547_{\pm 0.016}$	$\mathbf{0.822}_{\pm 0.011}$	$0.816_{\pm 0.011}$
Sepsis Prediction	4.8%	$0.621_{\pm 0.016}$	$0.667_{\pm 0.014}$	$0.707_{\pm 0.013}$	$\mathbf{0.786}_{\pm 0.012}$	$0.782_{\pm 0.013}$	$0.766_{\pm 0.013}$	$0.750_{\pm 0.014}$	$0.709_{\pm 0.015}$	$\mathbf{0.768}_{\pm 0.012}$	$0.763_{\pm 0.013}$

Steinberg et al. (2021), MOTOR Steinberg et al. (2024)), and dedicated survival models (CoxPH Cox (1972), DeepHit Lee et al. (2018), DeepSurv Katzman et al. (2018)). These models are trained under different objectives and expose different representational interfaces: XGBoost operates on engineered tabular features, CLMBR and MOTOR learn general-purpose representations from large longitudinal EHR corpora, survival models directly optimize time-to-event objectives, and EHR-WM is trained to predict action-conditioned latent trajectories. We therefore interpret this comparison as a benchmark of downstream predictive utility across modeling paradigms rather than a controlled architectural comparison.

The goal of this evaluation is not to demonstrate uniform superiority over specialized baselines. Instead, we ask whether representations forecasted through a trajectory-prediction objective retain sufficient information to support standard clinical prediction tasks while also exposing an explicit latent state trajectory for forecasting and action-conditioned simulation. The paired ORACLE/ROLLOUT evaluation further separates the predictive information available in the encoder-derived representation from degradation introduced during autoregressive rollout.

For representation-based methods, we use linear probes to assess how readily downstream information can be recovered from the learned representations Alain and Bengio (2016). For EHR-WM, the probe is evaluated independently at each valid latent state along the trajectory and performance is averaged across prediction timepoints; the same procedure is applied to both ORACLE and ROLLOUT states. CLMBR and MOTOR are evaluated using their latent representations.

Table 1 shows that performance is heterogeneous across tasks. EHR-WM is competitive with or exceeds the comparison models on several outcomes, including Acute MI, Hyperlipidemia, Hypertension, ICU admission, and Sepsis, while substantially stronger baselines remain on other tasks, including Lupus, Pancreatic Cancer, Long LOS, and 30-day readmission. We therefore do not interpret the results as evidence of uniform predictive superiority. Rather, they show that a model trained primarily to learn latent patient dynamics retains predictive information across a diverse collection of downstream tasks, with performance that depends on the endpoint and modeling paradigm.

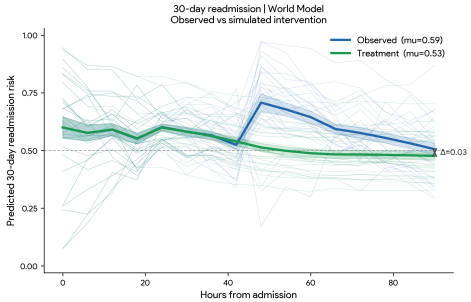

Figure 5: **Action-conditioned trajectory analysis for 30-day readmission.** Starting from the same initial state, EHR-WM generates different trajectories under the observed and alternative action sequences, illustrating the action sensitivity of the learned transition dynamics.

The comparison between ROLLOUT and ORACLE provides a separate view of rollout degradation. Across most tasks, their performance remains relatively close despite multi-step generation, indicating that the predictor preserves much of the task-relevant information encoded in the underlying latent trajectory. For highly imbalanced classification outcomes, we additionally report prevalence, AUPRC, and calibration analyses to complement AUROC in the Appendix.

Intervention-Conditioned Simulation The learned dynamics of EHR-WM support intervention-conditioned rollouts by propagating the same initial patient state under different clinical action sequences Verma et al. (2024). In Figure 5, we illustrate this behavior for 30-day readmission by comparing trajectories generated from the same initial state under the observed action sequence and an alternative action sequence. Both rollouts begin from the same latent patient state, such that differences in the resulting trajectories arise from changes to the action sequence supplied to the transition model. The alternative sequence produces a distinct time-resolved risk trajectory, demonstrating that the learned transition operator responds to changes in its clinical action input.

This experiment tests whether action conditioning has a functional effect on the learned dynamics. Holding the initial latent state fixed while changing the supplied action sequence produces corresponding changes in the generated trajectory. Because EHR-WM is trained on observational data, these differences demonstrate model sensitivity to action inputs and should not be interpreted as counterfactual or causal treatment-effect estimates.

Ablations Lastly, we conduct ablation studies to isolate the contributions of action conditioning and the training components used to stabilize long-horizon latent prediction. As shown in Figure 6, removing action conditioning degrades performance on several tasks, although the effect is task-dependent. More importantly, without conditioning on a_t, the transition model can only learn dynamics as a function of the patient state and its recent evolution, effectively averaging over variation associated with different observed treatment pathways. Conditioning F_θ on a_t instead allows the local transition dynamics to vary with the supplied clinical actions, providing the mechanism required for the intervention-conditioned rollout and perturbation analyses described above.

Figure 6: **Ablation of training dynamics and action conditioning.** Removing action conditioning reduces performance on selected tasks, while omitting teacher forcing, curriculum learning, or SIGReg leads to training collapse or unstable latent dynamics.

Model	ICU Adm.	Long LOS	Sepsis
Full Model	0.891	0.768	0.786
w/o Action Cond.	0.801	0.706	0.789
w/o Teacher Forcing	✗	✗	✗
w/o Curriculum	✗	✗	✗
w/o SIGReg	✗	✗	✗

Removing teacher forcing, curriculum learning, or SIGReg results in training collapse or unstable rollout behavior, consistent with known challenges in predictive world-model training Ennadir et al.. In noisy and partially observed clinical trajectories, training directly on long autoregressive rollouts exposes the model to compounding prediction error before reliable local transitions have been learned. Teacher forcing reduces this instability by periodically anchoring the rollout to encoder-derived latent states, while curriculum learning progressively increases the prediction horizon so that the model first learns short-range transitions before being exposed to longer free-running trajectories. SIGReg provides a complementary role by regularizing the geometry of the learned representation and discouraging degenerate latent solutions; related variance-preserving objectives such as VICReg have similarly been used to prevent representational collapse Bardes et al. (2021). In our ablations, removing these stabilization mechanisms causes the latent representation or rollout dynamics to degenerate, indicating that successful long-horizon training depends not only on the transition architecture but also on maintaining a well-conditioned predictive latent space.

5 Discussion

We formulate EHR modeling as learning action-conditioned latent patient dynamics rather than as a replacement for autoregressive EHR models. Across latent fidelity analyses, decoded forecasting, and downstream prediction, EHR-WM preserves useful predictive and temporal structure under multi-step rollout, although performance is task-dependent and specialized baselines remain stronger on several endpoints. The primary contribution is therefore an explicit trajectory interface that supports forecasting and action-sensitive simulation rather than uniform predictive superiority.

Limitations and Future Work Evaluation is restricted to MIMIC-IV and a high-frequency feature subset, limiting claims about cross-institutional and long-horizon generalization and partly explaining the smaller parameter footprint relative to foundation-model baselines. The multiplicative transition rule also limits action-driven displacement when $\Delta z_t \approx 0$; additive action terms or more flexible transition operators are natural extensions.

Ethics Statement. This work uses the publicly available, de-identified MIMIC-IV dataset and does not involve direct interaction with patients or collection of new human-subject data. Our models are developed for research purposes only and are not intended for clinical decision-making. In particular, action-conditioned rollouts characterize associations learned from observational data and should not be interpreted as causal treatment effects or treatment recommendations. Substantial external validation, causal evaluation, and prospective assessment would be required prior to any clinical deployment.

Reproducibility Statement. We provide detailed descriptions of the dataset construction, preprocessing, model architecture, optimization procedure, training schedule, hyperparameters, downstream tasks, and evaluation protocols in the main text and appendices. We additionally report the random seed, data splits, temporal discretization, feature-selection procedure, and computational setup used for training to facilitate reproduction of our results. We intend to release the full codebase upon acceptance of this work.

AI Use Statement. Generative AI tools were used solely to assist with polishing and improving the clarity of the manuscript's writing. They were not used to design the study, generate experimental results, perform data analysis, or determine the scientific conclusions. All AI-assisted text was reviewed and edited by the authors, who take full responsibility for the final manuscript.

REFERENCES

G. Alain and Y. Bengio. Understanding intermediate layers using linear classifier probes. *arXiv preprint arXiv:1610.01644*, 2016.

E. Alsentzer, J. Murphy, W. Boag, W.-H. Weng, D. Jindi, T. Naumann, and M. McDermott. Publicly available clinical bert embeddings. In *Proceedings of the 2nd clinical natural language processing workshop*, pages 72–78, 2019.

U. An, M. Jeong, S. A. Lee, A. Gorla, Y. Yang, and S. Sankararaman. Raptor: Scalable train-free embeddings for 3d medical volumes leveraging pretrained 2d foundation models. *Proceedings of machine learning research*, 267:1462, 2025a.

U. An, S. A. Lee, M. Jeong, A. Gorla, J. N. Chiang, and S. Sankararaman. Dk-behrt: Teaching language models international classification of disease (icd) codes using known disease descriptions. In *AAAI Bridge Program on AI for Medicine and Healthcare*, pages 133–143. PMLR, 2025b.

B. Arnrich, E. Choi, J. A. Fries, M. B. McDermott, J. Oh, T. Pollard, N. Shah, E. Steinberg, M. Wornow, and R. van de Water. Medical event data standard (meds): Facilitating machine learning for health. In *ICLR 2024 Workshop on Learning from Time Series For Health*, pages 03–08, 2024.

M. Assran, Q. Duval, I. Misra, P. Bojanowski, P. Vincent, M. Rabbat, Y. LeCun, and N. Ballas. Self-supervised learning from images with a joint-embedding predictive architecture. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 15619–15629, 2023.

M. Assran, A. Bardes, D. Fan, Q. Garrido, R. Howes, M. Muckley, A. Rizvi, C. Roberts, K. Sinha, A. Zholus, et al. V-jepa 2: Self-supervised video models enable understanding, prediction and planning. *arXiv preprint arXiv:2506.09985*, 2025.

R. Balestriero and Y. LeCun. Lejepa: Provable and scalable self-supervised learning without the heuristics. *arXiv preprint arXiv:2511.08544*, 2025.

A. Bardes, J. Ponce, and Y. LeCun. Vicreg: Variance-invariance-covariance regularization for self-supervised learning. *arXiv preprint arXiv:2105.04906*, 2021.

A. Bardes, J. Ponce, and Y. LeCun. Mc-jepa: A joint-embedding predictive architecture for self-supervised learning of motion and content features. *arXiv preprint arXiv:2307.12698*, 2023.

A. Bardes, Q. Garrido, J. Ponce, X. Chen, M. Rabbat, Y. LeCun, M. Assran, and N. Ballas. Revisiting feature prediction for learning visual representations from video. *arXiv preprint arXiv:2404.08471*, 2024.

Y. Bengio, J. Louradour, R. Collobert, and J. Weston. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 41–48, 2009.

S. Brotee, G. Chhetri, S. B. B. Polock, V. S. Bellamkonda, A. Rafe, and S. Das. A survey on joint embedding predictive architectures and world models. *Available at SSRN 5772122*, 2025.

Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu. Recurrent neural networks for multivariate time series with missing values, 2016. URL `https://arxiv.org/abs/1606.01865`.

R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, 2018.

T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

D. R. Cox. Regression models and life-tables. *Journal of the royal statistical society: Series B (methodological)*, 34(2):187–202, 1972.

E. De Brouwer, J. Simm, A. Arany, and Y. Moreau. Gru-ode-bayes: Continuous modeling of sporadically-observed time series. In *Advances in Neural Information Processing Systems*, 2019.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, 2019.

J. Ding, Y. Zhang, Y. Shang, Y. Zhang, Z. Zong, J. Feng, Y. Yuan, H. Su, N. Li, N. Sukiennik, et al. Understanding world or predicting future? a comprehensive survey of world models. *ACM Computing Surveys*, 58(3):1–38, 2025.

S. Ennadir, L. Zólyomi, and O. Smirnov. Understanding early collapse in predictive world-model pretraining. In *ICLR 2026 the 2nd Workshop on World Models: Understanding, Modelling and Scaling*.

R. S. Evans. Electronic health records: then, now, and in the future. *Yearbook of medical informatics*, 25(S 01):S48–S61, 2016.

Z. Fei, M. Fan, and J. Huang. A-jepa: Joint-embedding predictive architecture can listen. *arXiv preprint arXiv:2311.15830*, 2023.

A. Gorla and R. Puduppully. The illusion of generalization: Re-examining tabular language model evaluation. *arXiv e-prints*, pages arXiv–2602, 2026.

A. Gorla, R. Wang, Z. Liu, U. An, and S. Sankararaman. Cacti: leveraging copy masking and contextual information to improve tabular data imputation. *Proceedings of machine learning research*, 267:20187, 2025.

L. Guan, K. Valmeekam, S. Sreedharan, and S. Kambhampati. Leveraging pre-trained large language models to construct and utilize world models for model-based task planning. *Advances in Neural Information Processing Systems*, 36:79081–79094, 2023.

D. Ha and J. Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2(3):440, 2018.

Z. Jing, V. Jeanselme, Y. Kobayashi, S. A. Lee, C. Pang, A. Kashyap, Y. Li, X. Jiang, and S. Joshi. One loss to rule them all: Marked time-to-event for structured ehr foundation models. *arXiv preprint arXiv:2602.00541*, 2026.

A. E. Johnson, L. Bulgarelli, L. Shen, A. Gayles, A. Shammout, S. Horng, T. J. Pollard, S. Hao, B. Moody, B. Gow, et al. Mimic-iv, a freely accessible electronic health record dataset. *Scientific data*, 10(1):1, 2023.

J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger. Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network. *BMC medical research methodology*, 18(1):24, 2018.

P. Kidger, J. Morrill, J. Foster, and T. Lyons. Neural controlled differential equations for irregular time series. In *Advances in Neural Information Processing Systems*, 2020.

A. Kolo, C. Pang, E. Choi, E. Steinberg, H. Jeong, J. Gallifant, J. A. Fries, J. N. Chiang, J. Oh, J. Xu, et al. Meds decentralized, extensible validation (meds-dev) benchmark: Establishing reproducibility and comparability in ml for health. 2024.

A. M. Lamb, A. G. ALIAS PARTH GOYAL, Y. Zhang, S. Zhang, A. C. Courville, and Y. Bengio. Professor forcing: A new algorithm for training recurrent networks. *Advances in neural information processing systems*, 29, 2016.

A. Larey, E. Dahan, A. Bleiweiss, R. Kellerman, G. Leib, O. Nayshool, D. Ofer, T. Zinger, D. Dominissini, G. Rechavi, et al. Jepa-dna: Grounding genomic foundation models through joint-embedding predictive architectures. *arXiv preprint arXiv:2602.17162*, 2026.

Y. LeCun et al. A path towards autonomous machine intelligence version 0.9. 2, 2022-06-27. *Open Review*, 62(1):1–62, 2022.

C. Lee, W. Zame, J. Yoon, and M. Van Der Schaar. Deephit: A deep learning approach to survival analysis with competing risks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

S. A. Lee and K. Akamatsu. Foundation models for physiological signals: Opportunities and challenges, August 2025.

S. A. Lee, C. Tanade, H. Zhou, J. Lee, M. Thukral, B. Lu, and S. A. Desai. Towards on-device foundation models for raw wearable signals. In *NeurIPS 2025 Workshop on Learning from Time Series for Health*, 2025a.

S. A. Lee, A. Wu, and J. N. Chiang. Clinical modernbert: An efficient and long context encoder for biomedical text. *arXiv preprint arXiv:2504.03964*, 2025b.

S. A. Lee, C. Tanade, H. Zhou, J. Lee, M. Thukral, M. S. H. Khan, K. S. Chun, B. Lu, M. Gwak, M. B. Morshed, V. Nathan, M. M. Rahman, L. Zhu, S. Venkatraman, and S. A. Desai. HiMAE: Hierarchical masked autoencoders discover resolution-specific structure in wearable time series. In *The Fourteenth International Conference on Learning Representations*, 2026.

Y. Li, S. Rao, J. R. Ayala Solares, A. Hassaine, R. Ramakrishnan, D. Canoy, Y. Zhu, K. Rahimi, and G. Salimi-Khorshidi. Behrt: Transformer for electronic health records. *Scientific Reports*, 10:7155, 2020. doi: 10.1038/s41598-020-62922-y.

Z. Li, K. Natarajan, W. Zhang, M. Zhou, S. A. Lee, Y. Zhang, M. A. Xu, Z. Esmaeilpour, F. D. Salim, M. Malhotra, et al. Glucofm: A dual-stream foundation model for continuous glucose monitoring. *arXiv preprint arXiv:2605.30865*, 2026.

I. Loshchilov and F. Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

F.-M. Luo, T. Xu, H. Lai, X.-H. Chen, W. Zhang, and Y. Yu. A survey on model-based reinforcement learning. *Science China Information Sciences*, 67(2):121101, 2024.

Y. Luo, X. Chen, Y. Chen, L. Zhang, M. M. Islam, S. Zhao, P. Kotanko, S. Dasgupta, A. Campbell, R. Malhotra, et al. Toward world modeling of physiological signals with chaos-theoretic balancing and latent dynamics. *arXiv preprint arXiv:2605.15465*, 2026.

G. Lutsker, G. Sapir, J. Merino, S. Shilo, A. Godneva, E. Meirom, S. Mannor, H. Rossman, G. Chechik, and E. Segal. Simulating clinical interventions with a generative multimodal model of human physiology. *arXiv preprint arXiv:2604.27899*, 2026.

T. M. Moerland, J. Broekens, A. Plaat, and C. M. Jonker. Model-based reinforcement learning: A survey. *Foundations and Trends in Machine Learning*, 16(1):1–118, 2023.

J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang. Segment anything in medical images. *Nature communications*, 15(1):654, 2024.

L. Maes, Q. L. Lidec, D. Scieur, Y. LeCun, and R. Balestriero. Leworldmodel: Stable end-to-end joint-embedding predictive architecture from pixels. *arXiv preprint arXiv:2603.19312*, 2026.

M. B. McDermott, J. Xu, T. S. Bergamaschi, H. Jeong, S. A. Lee, N. Oufattole, P. Rockenschaub, K. Stankevičiūtė, E. Steinberg, J. Sun, et al. Meds: Building models and tools in a reproducible health ai ecosystem. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pages 6243–6244, 2025.

R. Miotto, L. Li, B. A. Kidd, and J. T. Dudley. Deep patient: An unsupervised representation to predict the future of patients from the electronic health records. *Scientific Reports*, 6:26094, 2016. doi: 10.1038/srep26094.

L. Mu, Z. Huang, Y. Gu, S. Qin, S. Zhang, and X. Zhang. Ehrworld: A patient-centric medical world model for long-horizon clinical trajectories. *arXiv preprint arXiv:2602.03569*, 2026.

H. Nam, Q. L. Lidec, L. Maes, Y. LeCun, and R. Balestriero. Causal-jepa: Learning world models through object-level latent interventions. *arXiv preprint arXiv:2602.11389*, 2026.

S. Narvekar, B. Peng, M. Leonetti, J. Sinapov, M. E. Taylor, and P. Stone. Curriculum learning for reinforcement learning domains: A framework and survey. *Journal of Machine Learning Research*, 21(181):1–50, 2020.

M. Odgaard, K. V. Klein, S. M. Thysen, E. Jimenez-Solem, M. Sillesen, and M. Nielsen. Core-behrt: A carefully optimized and rigorously evaluated behrt. *arXiv preprint arXiv:2404.15201*, 2024.

A. Oussidi and A. Elhassouny. Deep generative models: Survey. In *2018 International conference on intelligent systems and computer vision (ISCV)*, pages 1–8. IEEE, 2018.

Y. Pan, J. Sha, S. A. Lee, M. Cannesson, W. Wang, and J. N. Chiang. Hipno: Symmetry-aware physics-informed neural operators for noninvasive hemodynamic inference. *arXiv preprint arXiv:2608.10011*, 2026.

C. Pang, X. Jiang, K. S. Kalluri, M. Spotnitz, R. Chen, A. Perotte, and K. Natarajan. Cehr-bert: Incorporating temporal information from structured ehr data to improve prediction tasks. In *Proceedings of Machine Learning for Health*, volume 158 of *Proceedings of Machine Learning Research*, pages 239–260, 2021.

C. Pang, V. Jeanselme, Y. S. Choi, X. Jiang, Z. Jing, A. Kashyap, Y. Kobayashi, Y. Li, F. Pollet, K. Natarajan, et al. Fomoh: A clinically meaningful foundation model evaluation for structured electronic health records. *arXiv preprint arXiv:2505.16941*, 2025.

L. Rasmy, Y. Xiang, Z. Xie, C. Tao, and D. Zhi. Med-bert: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *npj Digital Medicine*, 4:86, 2021. doi: 10.1038/s41746-021-00455-y.

Y. Rubanova, R. T. Q. Chen, and D. Duvenaud. Latent odes for irregularly-sampled time series, 2019a. URL `https://arxiv.org/abs/1907.03907`.

Y. Rubanova, R. T. Q. Chen, and D. Duvenaud. Latent odes for irregularly-sampled time series. In *Advances in Neural Information Processing Systems*, 2019b.

D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. Technical report, 1985.

R. Salakhutdinov. Learning deep generative models. *Annual Review of Statistics and Its Application*, 2(1):361–385, 2015.

R. Sekar, O. Rybkin, K. Daniilidis, P. Abbeel, D. Hafner, and D. Pathak. Planning to explore via self-supervised world models. In *International conference on machine learning*, pages 8583–8592. PMLR, 2020.

A. Shmatko, A. W. Jung, K. Gaurav, S. Brunak, L. H. Mortensen, E. Birney, T. Fitzgerald, and M. Gerstung. Learning the natural history of human disease with generative transformers. *Nature*, 647(8088):248–256, 2025.

E. Steinberg, K. Jung, J. A. Fries, C. K. Corbin, S. R. Pfohl, and N. H. Shah. Language models are an effective representation learning technique for electronic health record data. *Journal of biomedical informatics*, 113:103637, 2021.

E. Steinberg, Y. Xu, J. A. Fries, and N. Shah. MOTOR: A time-to-event foundation model for structured medical records. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=NialiwI2V6.

H. Thimonier, J. L. D. M. Costa, F. Popineau, A. Rimmel, and B.-L. Doan. T-jepa: Augmentation-free self-supervised learning for tabular data. *arXiv preprint arXiv:2410.05016*, 2024.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need, 2023. URL https://arxiv.org/abs/1706.03762.

S. Verma, V. Boonsanong, M. Hoang, K. Hines, J. Dickerson, and C. Shah. Counterfactual explanations and algorithmic recourses for machine learning: A review. *ACM Computing Surveys*, 56(12): 1–42, 2024.

J. Wang, X. Chen, J. He, X. Xu, S. Xie, and F. Han. Chronomedicalworld: A medical world model for learning patient trajectories from longitudinal care data. *arXiv preprint arXiv:2605.21963*, 2026a.

Y. Wang, O. Bounou, G. Zhou, R. Balestriero, T. G. Rudner, Y. LeCun, and M. Ren. Temporal straightening for latent planning. *arXiv preprint arXiv:2603.12231*, 2026b.

M. Wornow, R. Thapa, E. Steinberg, J. A. Fries, and N. H. Shah. Ehrshot: An ehr benchmark for few-shot evaluation of foundation models. In *Advances in Neural Information Processing Systems (Datasets and Benchmarks Track)*, 2023a.

M. Wornow, Y. Xu, R. Thapa, B. Patel, E. Steinberg, S. Fleming, M. A. Pfeffer, J. Fries, and N. H. Shah. The shaky foundations of large language models and foundation models for electronic health records. *npj digital medicine*, 6(1):135, 2023b.

Y. Zhang and S. Li. Chronoformer: Time-aware transformer architectures for structured clinical event modeling. *arXiv preprint arXiv:2504.07373*, 2025.

H. Zhou, S. A. Lee, C. Tanade, K. S. Chun, J. Lee, M. Gwak, M. Thukral, J. Sung, E. Hwang, M. B. Morshed, L. Zhu, V. Nathan, M. M. Rahman, S. Venkatraman, and S. A. Desai. Physiology-aware masked cross-modal reconstruction for biosignal representation learning. *arXiv preprint arXiv:2605.00973*, 2026.

APPENDIX

A BROADER IMPACTS

Predictive world models may provide a useful representation-learning framework for longitudinal EHR by exposing explicit, action-conditioned latent trajectories rather than only static or pooled patient representations. In this work, alternative-action rollouts characterize sensitivity of learned observational dynamics and should not be interpreted as causal estimates or treatment recommendations. The model is evaluated only on MIMIC-IV and a restricted feature space, and substantial external validation, causal methodology, and prospective evaluation would be required before considering clinical decision support. The compact architecture may reduce computational requirements for research use, but parameter count alone does not establish deployment efficiency or clinical utility.

B EXTENDED RELATED WORKS

Neural differential equations and controlled dynamics: Neural differential equation models place continuous-time dynamics at the core of representation learning by parameterizing a vector field with a neural network and integrating it with a numerical solver. The canonical Neural ODE formulation treats depth (or time) as continuous, yielding a flexible continuous-time analogue of residual networks and enabling adaptive computation through solver tolerances Chen et al. (2018). For irregularly sampled sequences, this perspective has motivated latent-variable and hybrid formulations in which latent states evolve continuously between observations and are updated at observation times, allowing principled handling of non-uniform gaps and missingness Rubanova et al. (2019b); De Brouwer et al. (2019).

A key limitation of autonomous ODE dynamics in the EHR setting is semantic mismatch: an ODE trajectory is determined by its initial condition and the autonomous flow, so the latent state may drift even in the absence of new clinical evidence. Controlled differential equations address this by defining state evolution as a functional of a *driving path* (the control), so that state changes are induced by observed changes in the input signal. Neural Controlled Differential Equations (Neural CDEs) operationalize this idea for irregular multivariate time series by learning a state-conditioned linear map from infinitesimal control increments to infinitesimal state increments, yielding a model that is naturally aligned with event-driven data collection Kidger et al. (2020). Our work inherits this controlled-dynamics viewpoint but uses it to reframe EHR foundation modeling: rather than treating continuous-time modeling as a technical fix for irregular sampling, we treat it as the right object of representation learning, a population-level geometry of *health-state transitions*.

C DATASETS

Table 2: Dataset Split of MIMIC-IV

Split Name	Patients Count	Event Count	Number of Transitions
Training Set	239,770	424,391,148	6,370,510
Tuning Set	29,971	51,981,308	787,619
Test Set	29,971	52,537,879	791,916

Table 3: Patient demographics by data split (full MIMIC-IV)

Variable	**Train**	**Val**	**Test**
N patients	239,770	29,971	29,971
N encounters	345,348	42,836	43,047
Sex, n (%)			
Male	113,035 (47.1%)	14,038 (46.8%)	14,086 (47.0%)
Female	126,735 (52.9%)	15,933 (53.2%)	15,885 (53.0%)
Unknown	0 (0.0%)	0 (0.0%)	0 (0.0%)
Age at admission, years			
Mean (SD)	56.2 (20.2)	56.5 (20.2)	56.3 (20.2)
Median [IQR]	57.6 [39.5–72.0]	57.9 [39.7–72.4]	57.6 [39.9–72.2]
BMI (kg/m^2)			
Mean (SD)	28.4 (6.8)	28.3 (6.7)	28.4 (6.8)
Median [IQR]	27.2 [23.7–31.7]	27.2 [23.7–31.5]	27.2 [23.7–31.7]
Race / ethnicity, n (%)			
White	96,347 (40.2%)	12,043 (40.2%)	12,064 (40.3%)
Black / African American	17,779 (7.4%)	2,241 (7.5%)	2,170 (7.2%)
Hispanic / Latino	7,566 (3.2%)	907 (3.0%)	935 (3.1%)
Asian	5,471 (2.3%)	698 (2.3%)	709 (2.4%)
Other	6,531 (2.7%)	858 (2.9%)	914 (3.0%)
Unknown	106,076 (44.2%)	13,224 (44.1%)	13,179 (44.0%)

We construct longitudinal patient trajectories from MIMIC-IV Johnson et al. (2023) in the MEDS format McDermott et al. (2025); Arnrich et al. (2024) and partition patients into disjoint training, validation, and test splits at the patient level to prevent information leakage [2]. The resulting dataset consists of approximately 240K patients in the training set and 30K patients each in validation and test, corresponding to over 424M clinical events and 6.3M transitions in the training split. As shown in Table 3, demographic characteristics are well balanced across splits, with consistent distributions in sex, age, BMI, and race/ethnicity.

[2] https://github.com/Medical-Event-Data-Standard/MIMIC_IV_MEDS

D Downstream Tasks

Downstream tasks follow established EHRShot, FoMoH, and MEDS-DEV benchmark definitions (Wornow et al., 2023a; Pang et al., 2025; Kolo et al., 2024). We use them as standardized representation probes; acute-care outcomes are most directly aligned with the high-frequency inpatient setting, while longer-horizon disease tasks test whether the representation retains predictive information beyond immediate physiological dynamics.

D.1 Classification Evaluation

The goal of the classification evaluation is to predict future outcomes using only information available at each timepoint.

For disease incidence, we define a binary label over a 365-day horizon. A patient is positive if their first ICD code for a condition appears within one year after admission. Patients with no such code in that window are labeled negative. Those with a prior diagnosis before admission are excluded to ensure the model predicts new onset rather than existing disease.

We also include three operational outcomes tied to the hospital stay. Long length of stay is defined as a stay of at least 7 days. ICU transfer indicates whether any ICU-related event occurs during the admission. Thirty-day readmission indicates whether a new admission occurs within 30 days after discharge.

Patient representations are evaluated at valid prediction timepoints along the trajectory. For outcomes that occur during the observation period, only representations strictly preceding the outcome are included; representations at or after outcome occurrence are excluded from evaluation. For censored or negative examples, representations are included only while the patient remains at risk for the corresponding outcome. We apply the same evaluation protocol separately to embeddings derived from observed states and to embeddings obtained from predicted trajectories. For each task, we train a logistic regression classifier with standard regularization and balanced class weights. Models are evaluated on a held-out test set using AUROC with bootstrap confidence intervals. Patients with missing or inapplicable labels are excluded per task.

D.2 Time-to-Event (Survival) Evaluation

We evaluate time-to-event prediction using a standard right-censored survival setup. Each patient contributes a duration, an event indicator, and a representation. The duration measures time to the event or censoring, depending on whether the outcome occurs within the observation window.

For disease incidence, duration is time from admission to first diagnosis or censoring at one year. Patients without a diagnosis are censored. For long length of stay, the event is defined as remaining hospitalized for at least 7 days; patients who remain hospitalized to this threshold experience the event at day 7, whereas patients discharged before day 7 are censored at discharge. For ICU transfer, duration is time from admission to the first ICU-related event or discharge. For readmission, duration is time from discharge to the next admission, capped at 30 days, with censoring if no readmission occurs.

D.3 Forecasting

We evaluate forecasting capability via a linear decoding analysis on the learned patient embeddings. For each continuous target feature, we fit a probe on all training embeddings using only observed values, standardizing the embedding space before training. We then apply the same probe to held-out test embeddings and assess prediction quality separately across clinically meaningful time buckets, namely 0–6h, 6–12h, $(12–24)$h, and $(24–48)$h. This setup isolates how well the latent state supports downstream prediction at different stages of the patient stay, rather than conflating representation quality with temporal distribution shift. We evaluate the models using mean absolute error (MAE) and coefficient of determination R^2.

We compare against global mean, LOCF, Ridge regression, GRU-D Che et al. (2016), Latent ODE Rubanova et al. (2019a), and a Transformer forecaster Vaswani et al. (2017). All methods receive the same two-timestep initialization context and are evaluated on identical target variables, horizons, and

Table 4: EHR WM Encoder / projector architecture.

Block	Input shape	Output shape	Details
State input x_t	D_x	D_x	Concatenated clinical state features
Linear layer	D_x	256	Dense projection to hidden space
LayerNorm	256	256	Stabilizes feature scale
ReLU	256	256	Nonlinear feature transformation
Dropout	256	256	$p = 0.4$ regularization
Linear layer	256	128	Projection to latent space
Latent state z_t	128	128	Final representation

Table 5: EHR WM Predictor architecture

Block	Input shape	Output shape	Details
Latent state z_t	128	128	Current latent state
Action input a_t	d_a	d_a	Binary intervention vector
Concatenation	$128 + d_a$	$128 + d_a$	Joint state-action input
SpectralNorm Linear	$128 + d_a$	256	Lipschitz-constrained mapping
ReLU	256	256	Nonlinearity
SpectralNorm Linear	256	128×128	Transition operator parameters
Reshape	128×128	128×128	Matrix $F_\theta(z_t, a_t)$

held-out patients. This deliberately sparse context may disadvantage sequence models that benefit from longer observation histories, so results should be interpreted as short-context forecasting rather than each model's optimal-history performance.

To compare against nonlinear sequence models, we additionally evaluate GRU-D [3] (Che et al., 2016), which explicitly incorporates missingness and time since observation through learned decay mechanisms and is designed for sparsely observed multivariate clinical time series. Latent ODE [4](Rubanova et al., 2019a) represents the observed history in a latent state whose evolution is governed by a learned continuous-time ordinary differential equation, providing a continuous-dynamics comparison to our discrete action-conditioned transition model. Finally, we include a Transformer forecaster (Vaswani et al., 2017), which uses self-attention over the observed temporal context to model long-range dependencies and predict subsequent clinical measurements. All forecasting baselines receive the same input context and are evaluated over the same target variables, prediction windows, and held-out patients; performance is computed only at timepoints for which the corresponding ground-truth clinical variable is observed.

An important consideration in interpreting these comparisons is that forecasting is initialized from only the first two observed timepoints, which are required to initialize the EHR-WM dynamics through the latent increment Δz_t. This provides all models with a relatively short observation history and may disproportionately limit baselines designed to exploit longer temporal context. In particular, sequence models such as GRU-D and the Transformer, as well as Latent ODE models that infer dynamics from an observed trajectory, may benefit from richer histories over which temporal dependencies, missingness patterns, and patient-specific trends can be estimated. Consequently, performance in this setting should be interpreted as forecasting from a deliberately sparse initial context rather than as a comparison under each model's potentially optimal history length.

D.4 ACTION-SEQUENCE PERTURBATION ANALYSIS

Starting from a fixed initial latent state, we roll out EHR-WM under the observed action sequence and under alternative action sequences. A linear probe trained on representations from observed trajectories is kept fixed and applied to each rollout. Changes in probe output therefore measure the sensitivity of the learned latent dynamics to the supplied action sequence. This analysis does not identify causal treatment effects.

E MODEL OPTIMIZATION AND HYPERPARAMETERS

We provide a complete specification of architectural design, optimization procedure, and training hyperparameters used across all experiments to ensure reproducibility as seen in prior works. Unless otherwise stated, all configurations are held fixed across tasks to isolate representation quality from downstream model variation and ensure comparability between classification, survival, and forecasting evaluations.

The predictor outputs a matrix-valued function $F_\theta(z_t, a_t) \in \mathbb{R}^{D \times D}$, which defines a local linearization of the latent dynamics. Spectral normalization is applied to the predictor layers to regularize the Lipschitz behavior of the mapping from state-action inputs to transition-operator parameters and to improve stability during multi-step rollout. We separately characterize the spectral properties of the resulting matrices $F_\theta(z_t, a_t)$ empirically in Appendix G.1.

[3] https://github.com/fteufel/PyTorch-GRU-D

[4] https://docs.kidger.site/diffrax/examples/latent_ode/

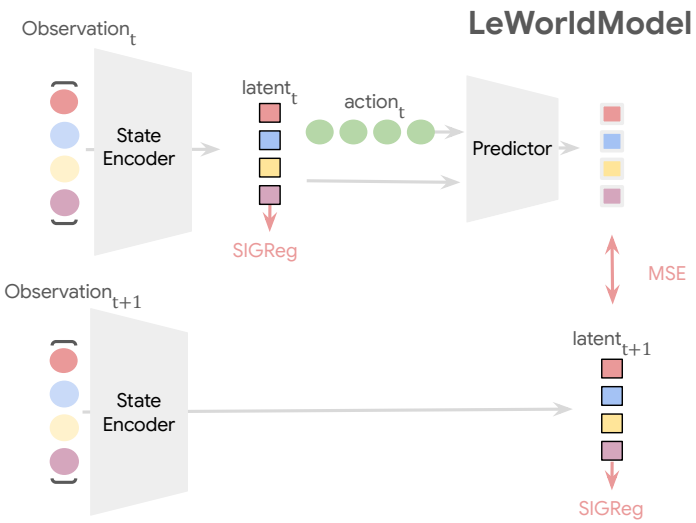

Figure 7: **Overview of the LeWorldModel framework.** At time t, observations are mapped into latent representations, which, together with actions, are passed to a predictor that models the subsequent latent state. The predicted latent is trained to align with the encoder-derived latent at time $t+1$ via a mean squared error objective, while structural regularization (SIGReg) is imposed on both current and future latent spaces to enforce consistency. This formulation enables learning latent dynamics directly in representation space without requiring reconstruction of the original observations.

E.1 Architecture

The model instantiates a controlled latent dynamical system in which observations $x_t \in \mathbb{R}^{D_x}$ are mapped to a compact latent state $z_t \in \mathbb{R}^D$, and temporal evolution is governed by a state- and action-dependent linear operator. The architecture decomposes into two components: (i) an encoder/projector that defines the latent representation, and (ii) a predictor that parameterizes the transition operator.

The encoder is implemented as a shallow projection network that maps high-dimensional, heterogeneous clinical inputs into a low-dimensional latent space. The use of LayerNorm and dropout induces stability and prevents overfitting in the presence of sparse and noisy inputs. The final latent dimensionality is fixed at $D = 128$, which we found to be sufficient to capture clinically meaningful variation while maintaining a compact state space.

E.2 Loss Function

Training is driven by a latent prediction objective over multi-step rollouts. Given predicted latent states $\hat{z}_{t+k}$ and encoded targets z_{t+k}, we minimize

$$\mathcal{L}_{\text{pred}} = \frac{1}{k} \sum_{j=1}^{k} \|\hat{z}_{t+j} - z_{t+j}\|_2^2.$$

To regularize the latent space, we incorporate a Sketched Isotropic Gaussian Regularization (SIGReg) term Balestriero and LeCun (2025), encouraging the empirical latent distribution to match $\mathcal{N}(0, I)$. The full objective is

$$\mathcal{L} = \lambda_{\text{mse}} \mathcal{L}_{\text{pred}} + \lambda_{\text{sigreg}} \mathcal{L}_{\text{SIGReg}}.$$

The prediction loss enforces temporal consistency in latent space, while the regularization term prevents collapse and ensures that representations remain well-conditioned for downstream linear probing.

Table 6: **Optimization hyperparameters.** Core training settings governing optimization dynamics, including learning rate, regularization, and gradient stabilization.

Parameter	Value
Optimizer	AdamW
Learning rate	3×10^{-4}
Weight decay	1×10^{-3}
Gradient clipping	1.0

Table 7: **Curriculum schedule.** Progressive rollout strategy controlling teacher forcing (ρ) and prediction horizon (k) across training stages.

Stage	ρ	k	Epochs	Patience
Stage 1	1.0	1	50	8
Stage 2	dynamic	2	50	16
Stage 3	dynamic	4	50	16
Stage 4	dynamic	8	50	16

Table 8: **Loss weighting coefficients.** Relative scaling of reconstruction and structural regularization objectives in the training loss.

Parameter	Value
λ_{mse}	1.0
λ_{sigreg}	0.05

Table 9: **Data and sequence configuration.** Key preprocessing and modeling choices defining sequence length, temporal resolution, and feature selection.

Parameter	Value
Sequence length	16
Batch size	64
Time bin size	6 hours
Top vitals	30
Top labs	50
Top medications	100
Minimum subject frequency	0.05
Random seed	42

E.3 Optimization and Training Schedule

We optimize all parameters using AdamW with decoupled weight decay Loshchilov and Hutter (2017). Let θ denote model parameters; updates take the form

$$\theta \leftarrow \theta - \eta \left(\nabla_\theta \mathcal{L} + \lambda_{\text{wd}} \theta \right),$$

with gradient clipping applied to enforce $\|\nabla_\theta\|_2 \leq 1.0$, stabilizing training under long rollout horizons.

Training follows a staged curriculum over rollout horizon k and teacher forcing ratio ρ Narvekar et al. (2020). Early stages use $\rho = 1$ (fully teacher-forced), ensuring accurate local dynamics. Subsequent stages progressively decrease dynamically ρ while increasing k, transitioning toward autoregressive multi-step prediction.

This curriculum can be interpreted as first learning accurate local Jacobian structure, followed by learning stable integration of the induced vector field over longer horizons.

E.4 Data and Sequence Configuration

Patient trajectories are discretized into fixed-length sequences of length $T = 16$, with each timestep representing a 6-hour aggregation window. This yields a regularly sampled approximation of the EHR event stream.

Feature selection is applied by restricting to the top-k most frequent variables within each modality (vitals, labs, medications), ensuring a consistent and dense representation across patients. Each continuous variable is represented by its normalized value and a binary observation mask. Unobserved values are set to zero after normalization with mask value zero, distinguishing missing observations from observed mean-valued measurements.

This discretization balances temporal fidelity with computational efficiency, while the restricted feature space mitigates sparsity and stabilizes optimization in the presence of heterogeneous clinical data.

E.5 Training Setup

EHR-WM was trained end-to-end on a single Apple M4 system with 48 GB unified memory using the MPS backend. Training required approximately 13 hours.

F Theoretical Motivation

F.1 Jacobian Induced Prior

Our predictor parameterizes a state- and action-dependent linear operator

$$F_\theta(z_t, a_t) \in \mathbb{R}^{D\times D},$$

which acts on latent increments and induces the update

$$z_{t+1} = z_t + F_\theta(z_t, a_t)\Delta z_t, \qquad \Delta z_t = z_t - z_{t-1}.$$

This choice can be interpreted as learning a local Jacobian field over latent space. Rather than predicting the next state directly, the model learns how infinitesimal changes in the current latent state propagate under interventions. In matrix form, a first-order local linearization of the latent transition map around (z_t, a_t) takes the form

$$z_{t+1} \approx z_t + J_G(z_t, a_t)\Delta z_t,$$

where

$$J_G(z_t, a_t) = \begin{bmatrix} \frac{\partial G_1}{\partial z_1} & \cdots & \frac{\partial G_1}{\partial z_D} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_D}{\partial z_1} & \cdots & \frac{\partial G_D}{\partial z_D} \end{bmatrix}$$

is the Jacobian of an underlying latent transition map G. Our model replaces this unknown Jacobian with a learned approximation $F_\theta(z_t, a_t)$.

The underlying clinical process is event-driven, while our model operates on its 6-hour discretization. We assume that adjacent latent states can often be described by local incremental changes and therefore use a Jacobian-inspired, state- and action-dependent operator to propagate finite latent increments. We interpret F_θ as a structured local parameterization of learned latent dynamics rather than an identified physiological Jacobian.

A simple stability argument follows from the operator norm of the learned velocity matrix. If

$$\|F_\theta(z_t, a_t)\|_2 \le \gamma$$

for all (z_t, a_t), then for any latent increment Δz_t,

$$\|z_{t+1} - z_t\|_2 = \|F_\theta(z_t, a_t)\Delta z_t\|_2 \le \gamma\|\Delta z_t\|_2.$$

Thus, when $\gamma < 1$, the propagated latent increment is contractive under a frozen-operator interpretation. Spectral normalization therefore biases the model toward damping increments during rollout; it does not imply contraction of the absolute or augmented state transition, which contains a residual identity component and additional state dependence.

The Jacobian formulation also gives a concrete interpretation of interventions. Since F_θ depends on a_t, two different action sequences induce different local linear operators, and hence different latent flows. This makes the model sensitive to treatment changes while remaining anchored to a single geometric object, namely the transition Jacobian field. As a result, the representation is both expressive and structured: it can represent heterogeneous clinical evolution, but only through locally coherent dynamical deformations.

The Jacobian-based parameterization provides three desirable properties for EHR modeling: local linear sensitivity to small changes, stability under multi-step rollout, and action-dependent modulation of latent dynamics. These properties make it a principled inductive bias for trajectory modeling in clinical time series.

F.2 Failure Mode: Absorbing State

Absorbing-state failure mode. If $\Delta z_t = 0$, then $\hat{z}_{t+1} = z_t$, producing a fixed point that is unresponsive to subsequent actions. SIGReg discourages global representational collapse, and we do not observe widespread fixed-point behavior empirically, but it does not remove this structural limitation. An additive action-forcing term is a natural extension.

G Extended Results

G.1 Jacobian and Stability Analysis

To better understand the learned latent dynamics, we analyze the spectrum and structure of the predictor-induced transition operators. Concretely, we extract latent trajectories from the held-out test set and evaluate the model in inference mode without gradient updates. For each patient sequence, we compute latent embeddings z_t using the encoder and construct latent increments $\Delta z_t = z_t - z_{t-1}$. At each timestep, we then evaluate the predictor to obtain the state- and action-conditioned velocity matrix

$$F_\theta(z_t, a_t) \in \mathbb{R}^{D \times D}.$$

This yields a large collection of velocity matrices, which we analyze using singular value decomposition and eigendecomposition. These matrices correspond to the learned local linearization of the latent dynamics and provide a direct view of how perturbations propagate under the model.

In addition to the velocity matrices, we compute the local state-transition Jacobian

$$J_t^{\text{local}} = \frac{\partial z_{t+1}}{\partial z_t}$$

via automatic differentiation on a random subset of test states. This quantity captures the local sensitivity of the next latent state to the current latent state, including both the explicit operator F_θ and its dependence on z_t. Because the recurrence also depends on z_{t-1}, this $D \times D$ quantity is a block of the full Jacobian of the augmented second-order system rather than the complete transition Jacobian. We derive the corresponding augmented Jacobian below. Together, the velocity operator F_θ and the local state-transition Jacobian characterize complementary aspects of the learned dynamics, including increment propagation, state dependence, conditioning, and local sensitivity.

Table 10: **Velocity matrix statistics for EHR-WM.** Summary of spectral and norm-based properties of the learned latent transition (velocity) operator, including measures of rank, conditioning, and magnitude, which together characterize the complexity, stability, and expressivity of the induced latent dynamics.

Metric	EHR-WM
Effective rank	106.68 ± 5.83
Condition number	4124.48 ± 46739.23
Frobenius norm	3.91 ± 0.60
Spectral norm	0.86 ± 0.14
Nuclear norm	38.50 ± 7.10

Velocity matrix statistics The learned velocity operators have high effective rank (106.68 of 128 dimensions) and mean spectral norm below one, indicating high-dimensional but damped increment dynamics. The large and variable condition number reflects strongly anisotropic propagation, with some latent directions attenuated substantially more than others.

Table 11: **Eigenvalue spectrum of F_θ.** Spectral analysis of the learned increment operator, reporting the distribution of eigenvalue magnitudes across sampled test states. Eigenvalues within the unit circle are consistent with locally damped dynamics under a frozen-operator interpretation.

Metric	EHR-WM
$\lvert\lambda\rvert < 1$	100.0%
$\lvert\lambda\rvert \in [0.9, 1.1]$	0.0%
$\lvert\lambda\rvert > 1$	0.0%
$\max \lvert\lambda\rvert$	0.938

Eigenvalue stability The eigenvalue spectrum provides a complementary view of the learned increment dynamics. Across the sampled test states, all eigenvalues of $F_\theta(z_t, a_t)$ lie strictly within the unit circle. Under a frozen-operator interpretation, this is consistent with locally damped increment dynamics rather than amplifying modes. Because $F_\theta(z_t, a_t)$ varies with both state and action during rollout, pointwise eigenvalue magnitudes below one do not by themselves establish global contraction or stability of the composed nonlinear dynamical system. We therefore interpret

this analysis as empirical evidence of locally stable increment propagation, complementary to the rollout-based stability results.

All sampled eigenvalues of F_θ lie within the unit circle, indicating contraction of the frozen increment dynamics. This does not imply contraction of the absolute state: because the update is second-order with a residual identity component, the augmented system is marginally stable under the frozen-coefficient approximation.

Full transition Jacobian. Because the update depends on both z_t and z_{t-1}, the model defines a second-order recurrence in the latent state. We therefore introduce the augmented state

$$x_t = \begin{bmatrix} z_t \\ z_{t-1} \end{bmatrix} \in \mathbb{R}^{2D},$$

and write the transition as

$$x_{t+1} = \Phi_t(x_t) = \begin{bmatrix} z_t + F_\theta(z_t, a_t)(z_t - z_{t-1}) \\ z_t \end{bmatrix}.$$

The corresponding full transition Jacobian is

$$J_t = \frac{\partial \Phi_t}{\partial x_t} = \begin{bmatrix} I + F_t + R_t & -F_t \\ I & 0 \end{bmatrix}, \qquad F_t := F_\theta(z_t, a_t),$$

where

$$R_t = \frac{\partial F_\theta(z_t, a_t)}{\partial z_t}(z_t - z_{t-1})$$

captures the additional state dependence of the learned transition operator. In contrast, the quantity

$$\frac{\partial z_{t+1}}{\partial z_t} = I + F_t + R_t$$

is only the $D \times D$ upper-left block of the full augmented Jacobian and should be interpreted as the local sensitivity of the next latent state to the current latent state, rather than as the complete Jacobian of the second-order dynamical system.

To characterize both components of the learned dynamics, we analyze the state-dependent operator F_t and the local state-transition block $\partial z_{t+1}/\partial z_t$ separately. The former describes how latent increments are propagated under the supplied clinical action, while the latter additionally incorporates the residual identity term and the dependence of F_θ on the current state. These quantities provide complementary views of the geometry and stability of the learned transition dynamics.

The local state-transition Jacobian

$$\frac{\partial z_{t+1}}{\partial z_t} = I + F_t + R_t$$

is nearly full-rank, with an effective rank close to the latent dimensionality $D = 128$. This indicates that the local mapping from the current latent state to the next latent state preserves variation across most latent directions rather than collapsing onto a substantially lower-dimensional subspace. Its spectral norm can exceed one because the transition contains a residual identity component, so local expansion in some directions is possible even when the increment operator F_t is damped. We therefore interpret these statistics as characterizing the local sensitivity and conditioning of the state transition, rather than as establishing global contraction of the full dynamical system.

Table 12: **Local state-transition Jacobian statistics.** Aggregate properties of the local state-transition Jacobian $\partial z_{t+1}/\partial z_t$, capturing its effective dimensionality, conditioning, and local sensitivity across sampled test states.

Metric	EHR-WM
Effective rank	127.00 ± 1.12
Condition number	4.16 ± 11.25
Spectral norm	1.24 ± 0.12

Spectral concentration The cumulative singular value spectrum reveals moderate concentration of energy: approximately one-third of the singular directions explain half of the operator norm, while over 100 directions are required to capture near-total energy. This indicates that the learned dynamics are neither low-rank nor fully isotropic, but instead exhibit structured high-dimensional behavior.

These results reveal a consistent picture of the learned dynamics. The velocity matrices exhibit high effective rank and predominantly damped local increment dynamics while maintaining substantial directional diversity. The local state-transition Jacobian remains nearly full-rank, indicating that the mapping from the current latent state to the next latent state preserves variation across most latent dimensions. Together with the empirical rollout analyses, these results suggest that the learned transition dynamics remain well-conditioned under multi-step prediction without requiring global contraction of the nonlinear dynamical system.

Table 13: **Rank for cumulative energy of F_θ.** Number of singular components required to capture increasing fractions of spectral energy, illustrating how variance is distributed across modes and indicating the degree of low-rank structure in the learned transition operator.

Threshold	EHR-WM
50%	40
75%	68
90%	92
95%	102
99%	116

G.2 GAN Discriminator over time

To complement the per-timestep analysis in Figure 8, we additionally evaluate discriminability at the trajectory level. Concretely, we construct a sequence-level classification task in which a discriminator D_ψ is trained to distinguish between encoder-derived latent trajectories $\{z_t\}_{t=1}^T$ and predictor-generated trajectories $\{\hat{z}_t\}_{t=1}^T$.

In parallel, we consider a stronger adversarial formulation in which D_ψ is optimized in a GAN-style setup to maximally separate the two trajectory distributions, effectively probing whether any systematic statistical signal remains that could distinguish generated rollouts from true latent dynamics. Notably, across both settings, discriminator performance remains consistently near chance (AUROC ≈ 0.5), even when evaluated per timestep and over full sequences. This suggests that the learned predictor induces a latent transition operator whose induced trajectory distribution closely matches that of the encoder, not only in terms of pointwise alignment but also in its temporal structure. These results indicate that rollout trajectories preserve distributional and temporal properties of the encoder-derived latent trajectories.

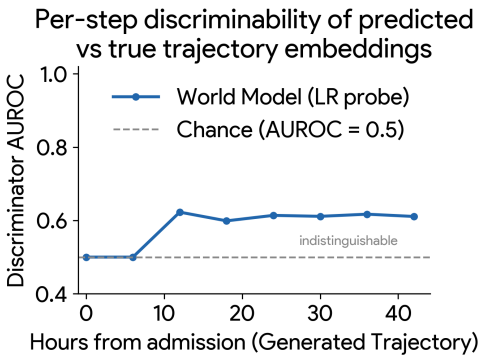

Figure 8: **Per-timestep discriminability of generated and encoder-derived latent states.** A linear discriminator is trained independently at each rollout timestep to distinguish predictor-generated from encoder-derived latent representations. Performance remains near chance (AUROC ≈ 0.5) across rollout depth, indicating limited linear separability between generated and encoder-derived states. Sequence-level discrimination is evaluated separately in the accompanying analysis.

G.3 Sensitivity to Temporal Discretization

We ablate the temporal bin size Δt to systematically characterize how discretization impacts representation learning and downstream predictive performance. Concretely, we evaluate four resolutions—3, 6, 9, and 12 hour bins—while holding the model architecture, optimization procedure, and training data fixed. This isolates the effect of temporal aggregation as a structural inductive bias on the learned latent dynamics. As shown in Table 14, performance consistently peaks at $\Delta t = 6$ hours across all evaluated tasks, indicating that this resolution achieves a favorable balance between temporal fidelity and statistical efficiency.

The observed trend reflects a trade-off that can be understood through the lens of signal-to-noise and effective sample complexity. At finer granularity ($\Delta t = 3$ hours), the sequence length increases, but each bin contains fewer observed events, leading to sparsity in the input space and higher variance in the resulting representations. This fragmentation reduces the reliability of both the encoder and the learned transition dynamics, particularly for infrequent clinical measurements. Conversely, coarser discretizations ($\Delta t = 9$ and 12 hours) aggregate over longer intervals, effectively smoothing the underlying trajectory and attenuating short-horizon variations that may carry predictive signal. This induces a bias toward overly stable dynamics and limits the model's ability to capture rapid physiological changes.

Table 14: **Ablation of temporal discretization**. Performance is best at $\Delta t = 6$ hours, reflecting a trade-off between sparsity at finer bins and loss of granularity at coarser bins.

Δt (hours)	ICU Adm.	Long LOS	Sepsis
3	0.882	0.760	0.779
6	0.891	0.768	0.786
9	0.884	0.761	0.781
12	0.876	0.754	0.774

G.4 Classification Performance

The AUPRC results in Table 15 show heterogeneous performance across tasks, with EHR-WM remaining competitive with the comparison models on several outcomes while underperforming on others. EHR-WM achieves the highest rollout AUPRC for Acute MI, ICU Admission, and Sepsis Prediction, and is also competitive for Hyperlipidemia and Hypertension. In contrast, MOTOR performs best on Lupus, Pancreatic Cancer, and Long LOS, while CLMBR achieves the strongest

Table 15: **AUPRC comparison across binary classification tasks with prevalence and WM rollout/oracle decomposition.** EHR-WM rollout denotes evaluation using predicted latent trajectories, while EHR-WM oracle denotes evaluation using encoder-derived latent trajectories from observed patient states.

Task	Prev	Binary Classification (AUPRC [Approximate 95% Interval])				
		GBM	CLMBR	MOTOR	EHR WM	
					Rollout	Oracle
Param Size		–	141M	143M	4.39M	
Acute MI	4.9%	0.191 [0.168–0.221]	0.194 [0.172–0.223]	0.208 [0.185–0.238]	0.214 [0.192–0.243]	0.205 [0.186–0.233]
Lupus	0.4%	0.015 [0.011–0.023]	0.019 [0.015–0.027]	0.022 [0.018–0.029]	0.009 [0.006–0.015]	0.008 [0.006–0.012]
Hyperlipidemia	30.9%	0.244 [0.225–0.260]	0.250 [0.231–0.265]	0.252 [0.233–0.268]	0.258 [0.240–0.273]	0.270 [0.252–0.284]
Hypertension	45.2%	0.204 [0.192–0.218]	0.271 [0.259–0.285]	0.278 [0.266–0.292]	0.291 [0.279–0.304]	0.294 [0.281–0.306]
Celiac	0.5%	0.023 [0.018–0.037]	0.009 [0.005–0.018]	0.012 [0.008–0.022]	0.012 [0.008–0.022]	0.013 [0.008–0.021]
Pancreatic Cancer	0.8%	0.082 [0.065–0.113]	0.081 [0.064–0.111]	0.102 [0.084–0.134]	0.043 [0.029–0.068]	0.049 [0.033–0.078]
ICU Admission	20.9%	0.325 [0.307–0.341]	0.314 [0.297–0.331]	0.368 [0.350–0.384]	0.410 [0.393–0.426]	0.410 [0.395–0.424]
Long LOS	15.0%	0.552 [0.532–0.572]	0.605 [0.585–0.624]	0.625 [0.606–0.645]	0.538 [0.518–0.558]	0.500 [0.480–0.520]
30-day Readmission	13.2%	0.318 [0.305–0.335]	0.413 [0.399–0.430]	0.353 [0.340–0.369]	0.226 [0.214–0.241]	0.234 [0.220–0.249]
Sepsis Prediction	4.8%	0.136 [0.109–0.163]	0.167 [0.139–0.195]	0.202 [0.174–0.232]	0.307 [0.275–0.340]	0.321 [0.291–0.354]

result for 30-day Readmission. Performance on Celiac is similar across several models, with all methods operating in a very low-prevalence regime.

The comparison between EHR-WM ROLLOUT and ORACLE representations is generally close across tasks, with only modest differences in AUPRC. The two are effectively identical for ICU Admission and remain similar for Acute MI, Lupus, Hypertension, Celiac, Pancreatic Cancer, and 30-day Readmission. Some larger differences are observed for Hyperlipidemia, Long LOS, and Sepsis Prediction, where the oracle representation performs differently from the rollout representation. Overall, the AUPRC evaluation complements the AUROC results by showing how predictive performance varies under substantial class imbalance while preserving the same task-dependent pattern across modeling approaches.

G.5 POST-HOC CALIBRATION

We further evaluate whether confidence estimates derived from EHR-WM representations remain calibrated as predictions are propagated forward in time. Calibration is performed post hoc on the downstream binary classifiers rather than on the world model itself. For each downstream task and rollout horizon, we first obtain the uncalibrated decision value s from the corresponding linear classifier and fit a Platt-scaling model on the held-out validation set,

$$\hat{p}(y = 1 \mid s) = \sigma(\alpha s + \beta),$$

where $\sigma(\cdot)$ denotes the logistic sigmoid and α, β are learned exclusively from validation examples. The fitted calibration mapping is then frozen and applied unchanged to the held-out test set. This procedure leaves both the learned world model and the downstream classifier fixed, and only learns a monotonic transformation from raw classifier scores to calibrated probabilities.

We assess calibration by comparing predicted probabilities with empirical outcome frequencies on the test set. For a perfectly calibrated classifier, examples assigned probability q should exhibit the positive outcome with empirical frequency approximately q. We summarize this relationship using reliability diagrams and expected calibration error (ECE). Specifically, predicted probabilities are partitioned into probability bins, and ECE is computed as

$$\mathrm{ECE} = \sum_{b=1}^{B} \frac{|S_b|}{N} \left|\mathrm{acc}(S_b) - \mathrm{conf}(S_b)\right|,$$

where S_b denotes the set of examples assigned to bin b, $\mathrm{acc}(S_b)$ is the empirical positive-outcome frequency within the bin, and $\mathrm{conf}(S_b)$ is the mean calibrated probability within the bin.

The first analysis compares reliability across selected rollout horizons. We report a one-step re-anchored setting as a short-horizon reference and compare it with free-running rollouts at increasing horizons k. This directly tests whether downstream risk estimates remain probabilistically calibrated as the latent trajectory is propagated progressively further from observed patient states. Deviations from the diagonal therefore reflect mismatch between predicted risk and empirical event frequency, rather than error in the underlying latent representation alone.

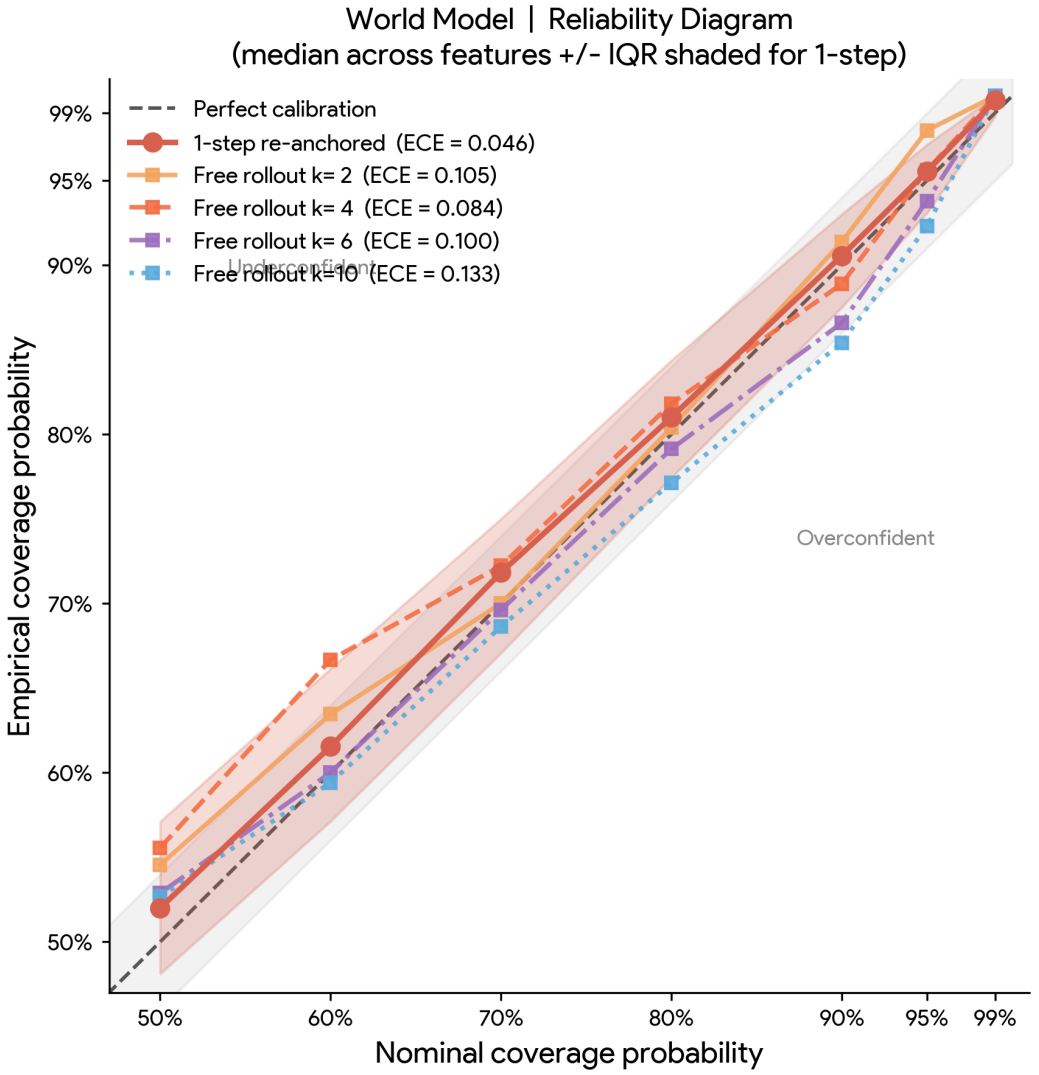

Figure 9: **Reliability of downstream binary predictions across rollout horizons.** Reliability diagram comparing Platt-scaled predicted probabilities with empirical positive-outcome frequencies for one-step re-anchored predictions and free-running rollouts at increasing horizons. The dashed diagonal denotes perfect calibration, while expected calibration error (ECE) summarizes the discrepancy between predicted confidence and empirical outcome frequency.

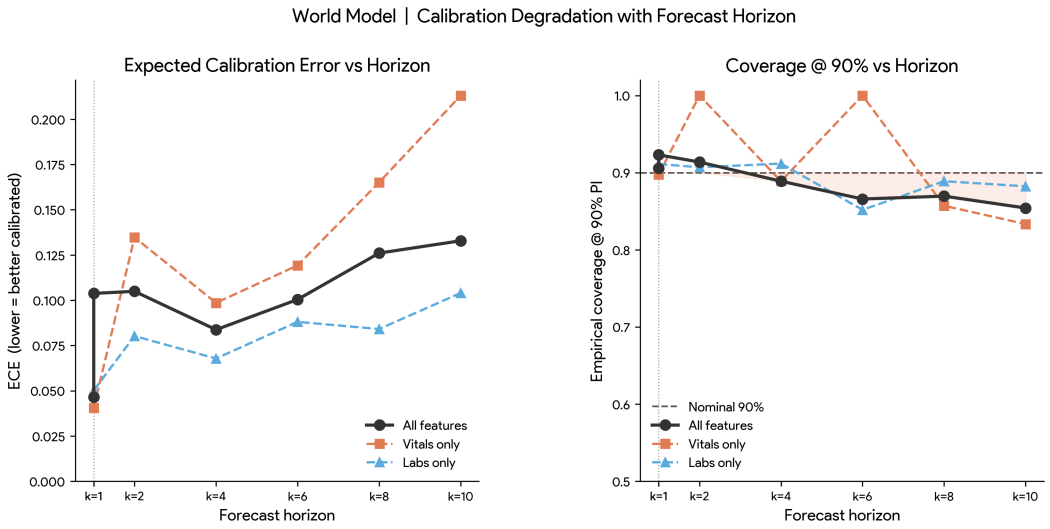

Figure 10: **Calibration degradation with increasing rollout horizon.** Expected calibration error (ECE) for downstream binary classifiers is reported across rollout horizons. Calibration is performed using Platt scaling fit on held-out validation data and applied unchanged to the test set. Increasing rollout depth generally produces larger calibration error, characterizing degradation in the probabilistic reliability of downstream predictions as the model moves further from observed patient states.

Across rollout horizons, the reliability curves remain broadly aligned with the diagonal, although calibration error generally increases with rollout depth. The one-step re-anchored setting provides the strongest calibration reference, whereas longer free-running rollouts expose the additional miscalibration introduced when downstream predictions are made from increasingly model-generated latent states.

The second analysis summarizes calibration explicitly as a function of rollout horizon. We report ECE across rollout lengths $k = 1, \ldots, 10$ for the evaluated binary downstream tasks. This allows us to quantify how probabilistic calibration changes as autoregressive rollout error accumulates. Because the same downstream classifier and validation-fitted Platt-scaling transformation are applied without refitting on the test set, differences across rollout horizons reflect changes in the quality of the generated latent representations rather than horizon-specific retraining.

Together, these analyses complement discrimination metrics such as AUROC and AUPRC. AUROC and AUPRC characterize whether the representation preserves ranking and class-separation information, whereas calibration evaluates whether the resulting predicted probabilities correspond to observed outcome frequencies. The combination therefore tests not only whether EHR-WM rollouts preserve task-relevant predictive information, but also whether confidence in downstream predictions remains meaningful as rollout depth increases.

G.6 Odds Ratios and Hazard Ratios by Demographic Groups

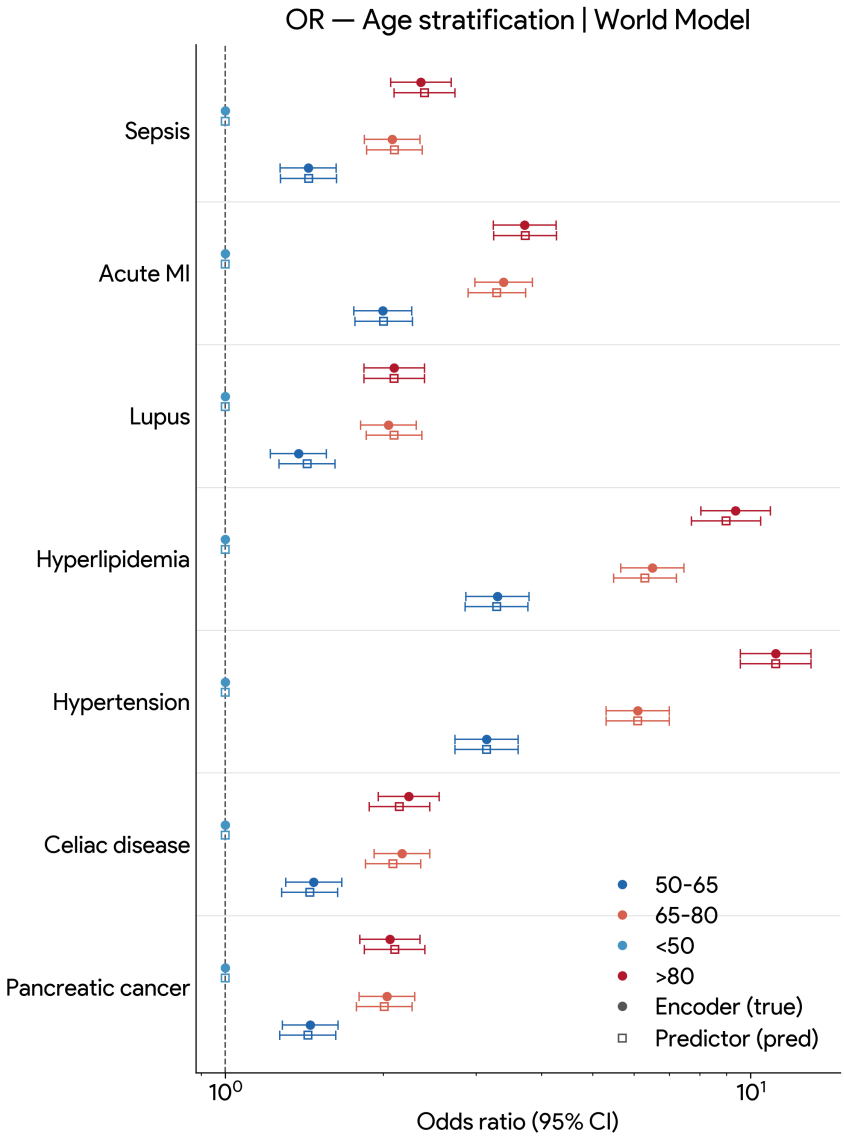

Figure 11: **Age-stratified odds ratios from generated trajectories.** Odds ratios (95% CI) computed from predictor-generated trajectories closely match those from encoder-derived representations across age groups, preserving relative risk structure and clinically consistent trends.

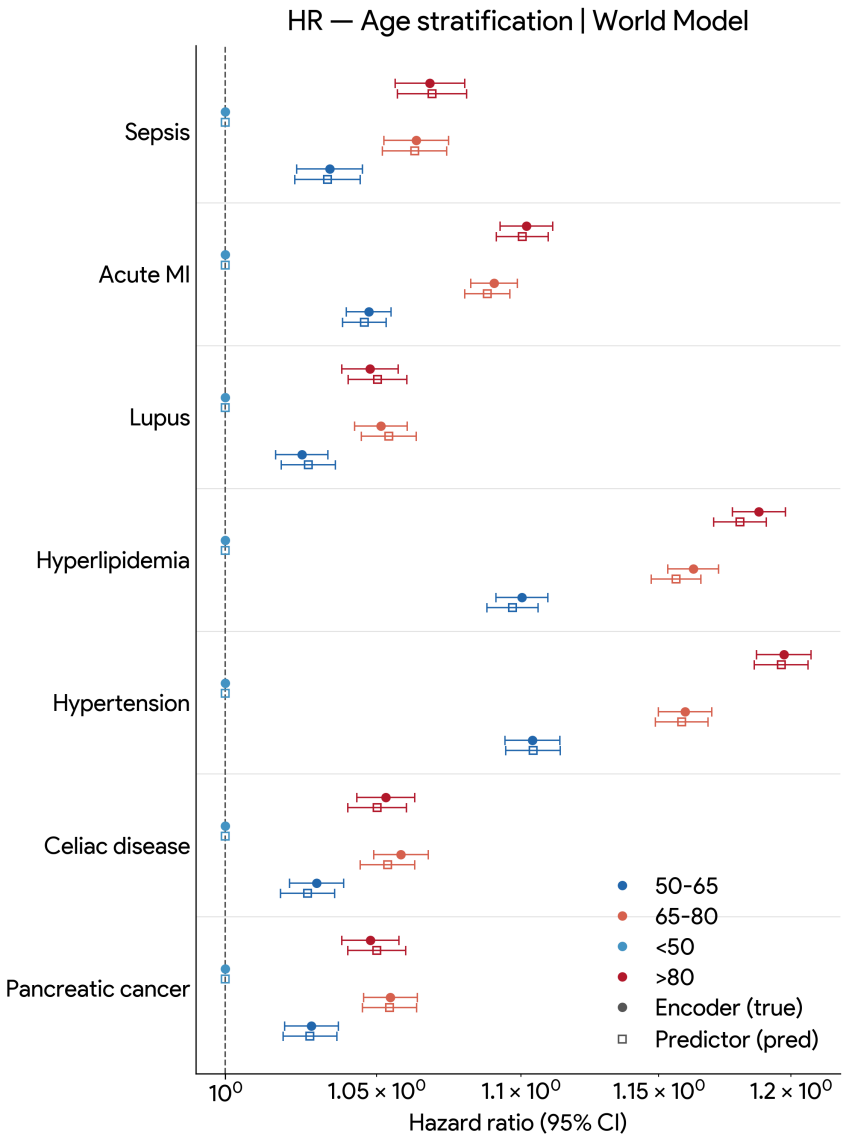

Figure 12: **Age-stratified hazard ratios from generated trajectories.** Hazard ratios (95% CI) computed from predictor-generated trajectories closely match those from encoder-derived representations across age groups, preserving relative risk structure and clinically consistent trends.

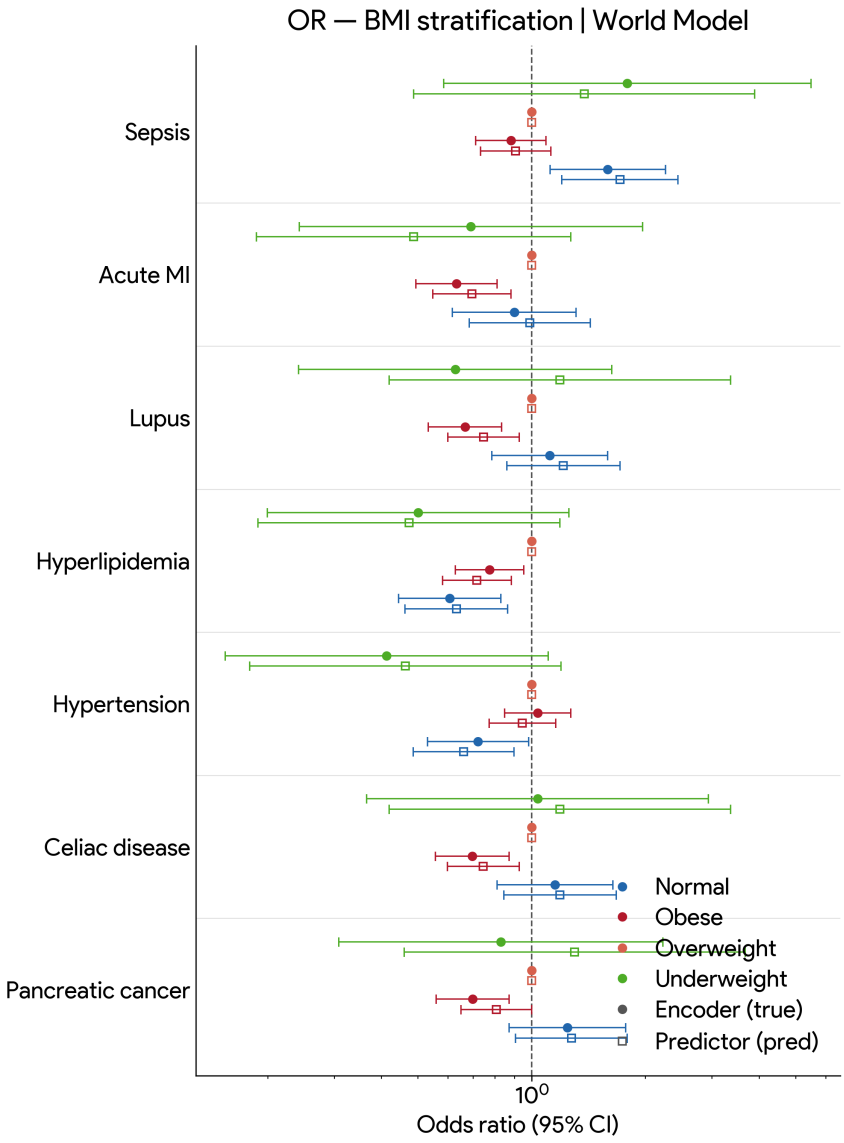

Figure 13: **BMI-stratified odds ratios from generated trajectories.** Odds ratios (95% CI) computed from predictor-generated trajectories closely match those from encoder-derived representations across BMI groups, preserving relative risk structure and clinically consistent trends.

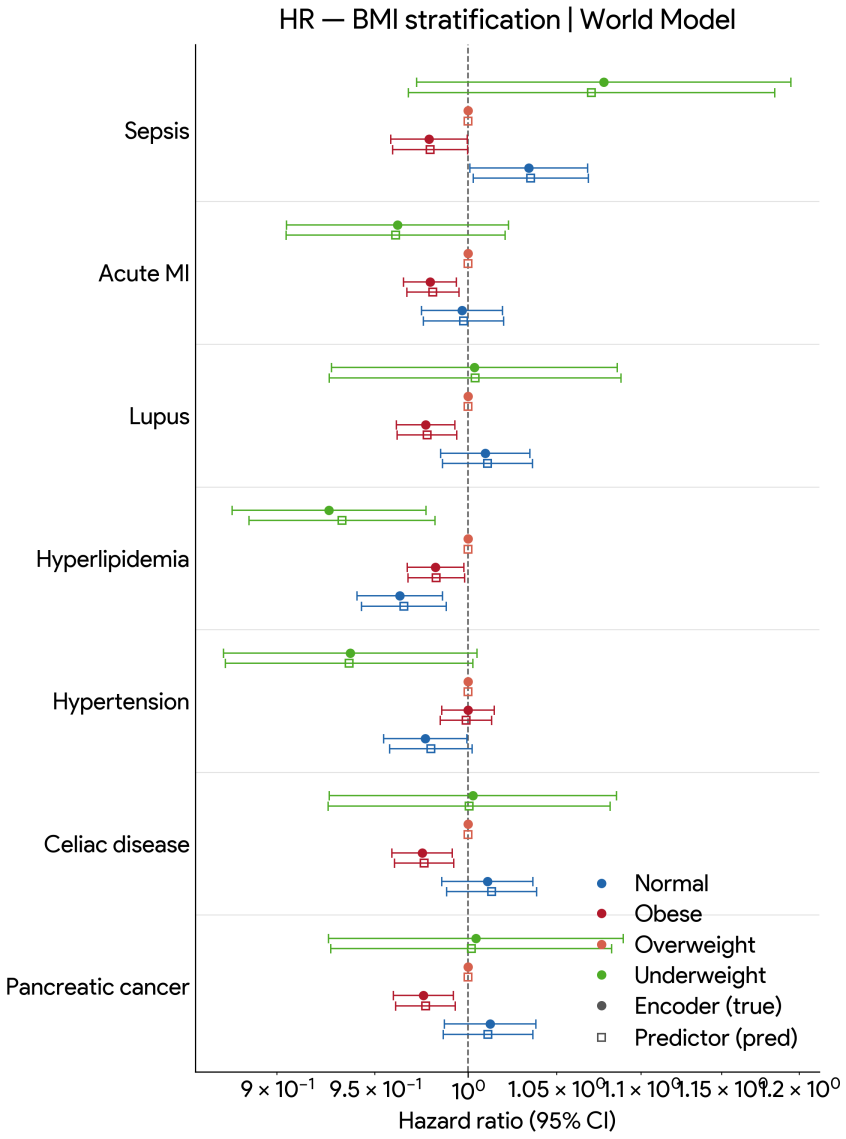

Figure 14: **BMI-stratified hazard ratios from generated trajectories.** Hazard ratios (95% CI) computed from predictor-generated trajectories closely match those from encoder-derived representations across BMI groups, preserving relative risk structure and clinically consistent trends.

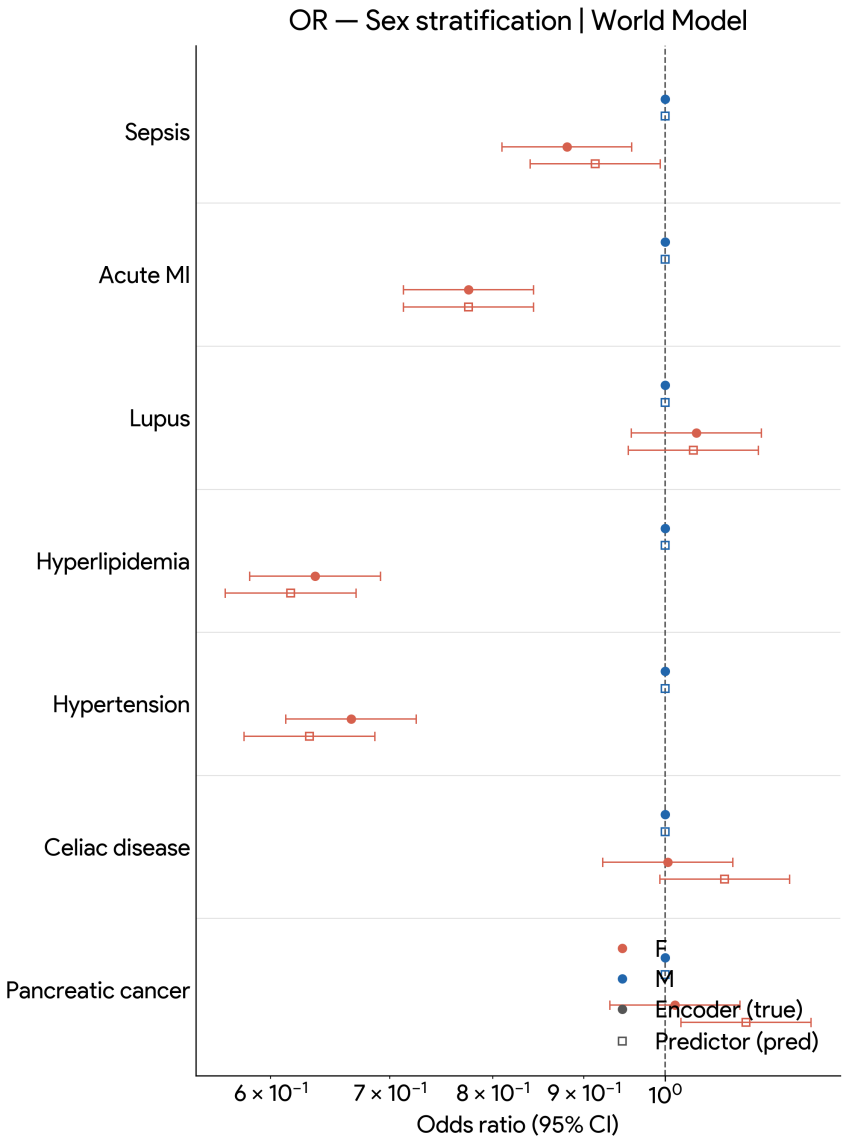

Figure 15: **Sex-stratified odds ratios from generated trajectories.** Odds ratios (95% CI) computed from predictor-generated trajectories closely match those from encoder-derived representations across sex groups, preserving relative risk structure and clinically consistent trends.

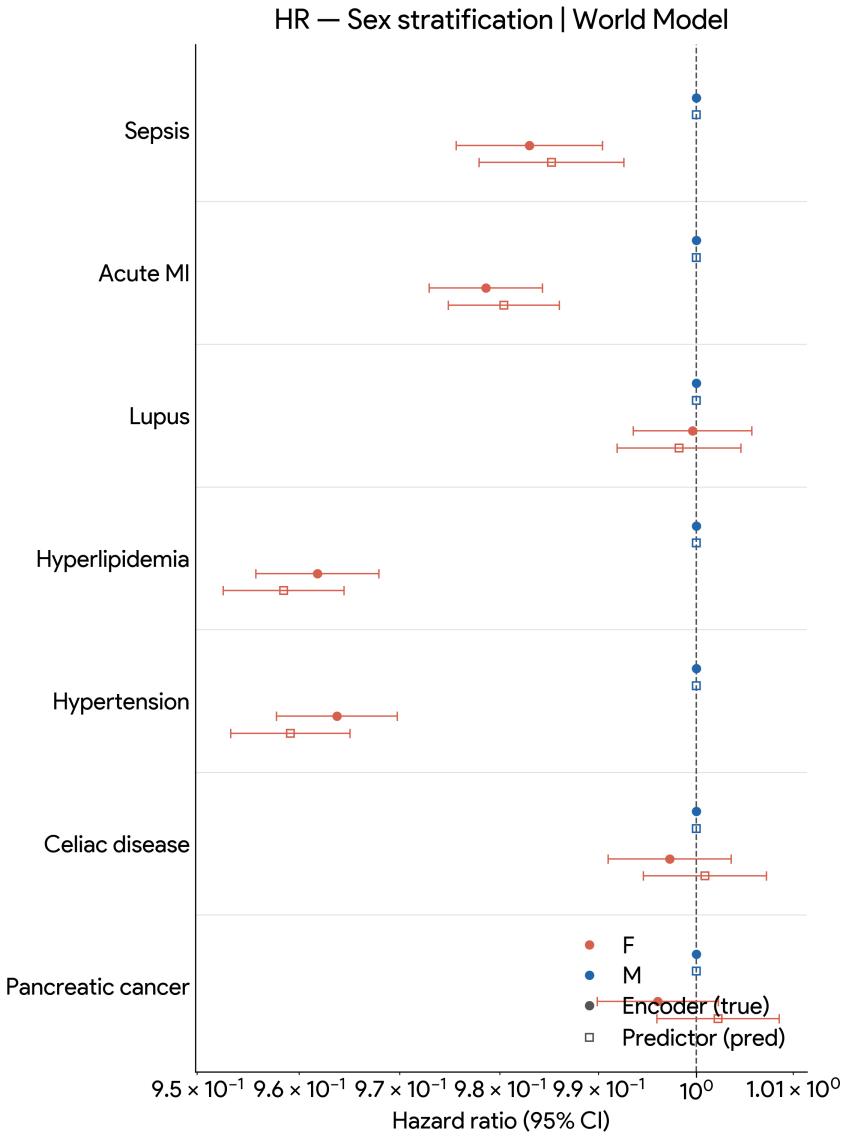

Figure 16: **Sex-stratified hazard ratios from generated trajectories.** Hazard ratios (95% CI) computed from predictor-generated trajectories closely match those from encoder-derived representations across sex groups, preserving relative risk structure and clinically consistent trends.

G.7 TEMPORAL PROBING

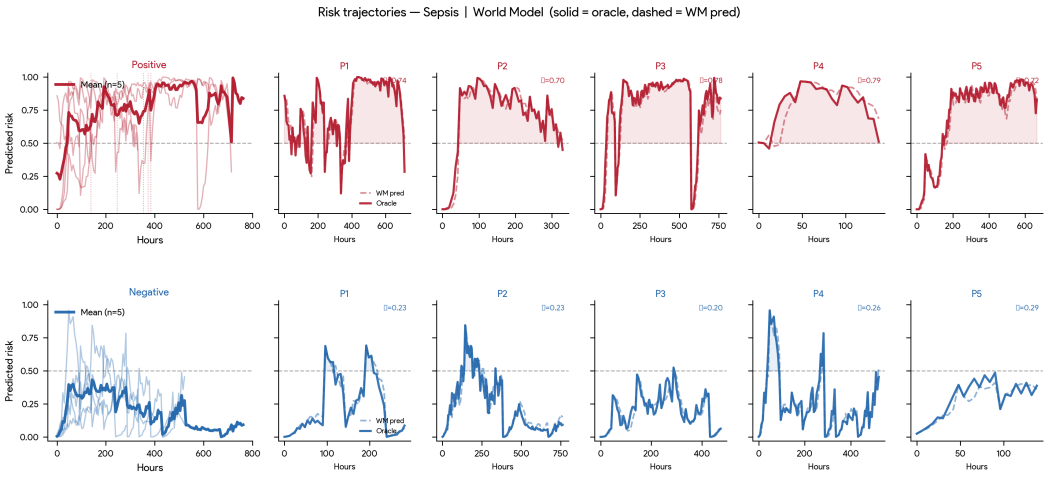

Figure 17: **Trajectory-level risk evolution under world model rollouts.** Predicted risk trajectories for sepsis comparing encoder-derived ("oracle", solid) and world model rollout ("WM pred", dashed) representations. Top row shows positive cases, bottom row negative cases, with cohort means (left) and representative patients (P1–P5). The model closely tracks temporal risk dynamics, capturing both sharp transitions and longer-term trends, while maintaining consistency with oracle trajectories across diverse clinical patterns.

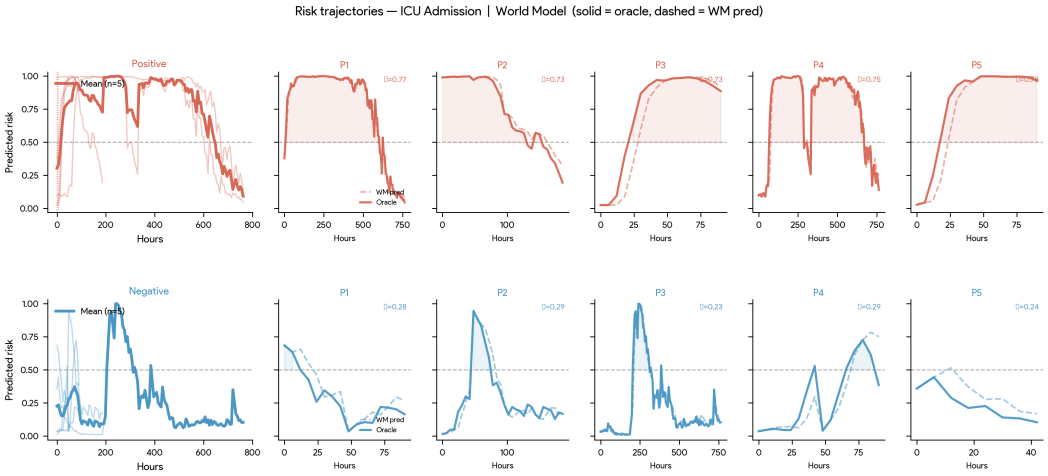

Figure 18: **Trajectory-level risk evolution under world model rollouts.** Predicted risk trajectories for ICU Admission comparing encoder-derived ("oracle", solid) and world model rollout ("WM pred", dashed) representations. Top row shows positive cases, bottom row negative cases, with cohort means (left) and representative patients (P1–P5). The model closely tracks temporal risk dynamics, capturing both sharp transitions and longer-term trends, while maintaining consistency with oracle trajectories across diverse clinical patterns.

G.8 FORECASTING MAE AND R^2



Figure 19: **Regression decoding performance across clinical variables.** Per-feature evaluation of forecasting accuracy comparing the world model (WM) against global mean and last observation carried forward (LOCF) baselines. Left: coefficient of determination (R^2); right: mean absolute error (MAE). Each point corresponds to a clinical variable. The world model consistently achieves higher R^2 and lower MAE across most features, indicating improved fidelity in capturing temporal dynamics and reconstructing future observations relative to static or persistence-based baselines.

G.9 TIME WINDOW REGRESSION

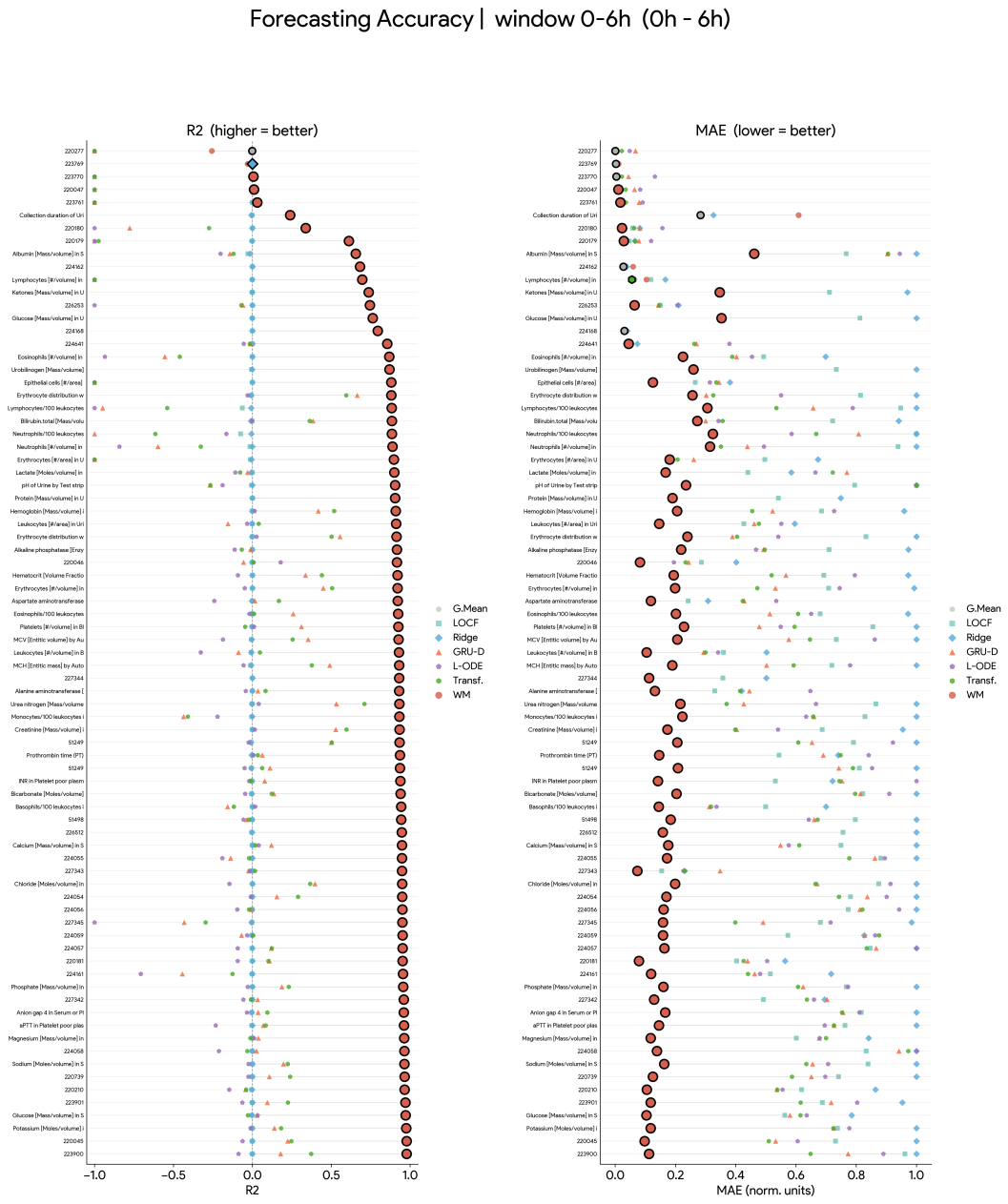

Figure 20: **Regression decoding performance at long-horizon rollout (0–6h).** Per-feature evaluation of forecasting accuracy for the 0–6 hour prediction window, with features sorted by world model (WM) R^2 in descending order. Left: coefficient of determination (R^2); right: mean absolute error (MAE).



Figure 21: **Regression decoding performance at long-horizon rollout (6–12h).** Per-feature evaluation of forecasting accuracy for the 6–12 hour prediction window, with features sorted by world model (WM) R^2 in descending order. Left: coefficient of determination (R^2); right: mean absolute error (MAE).

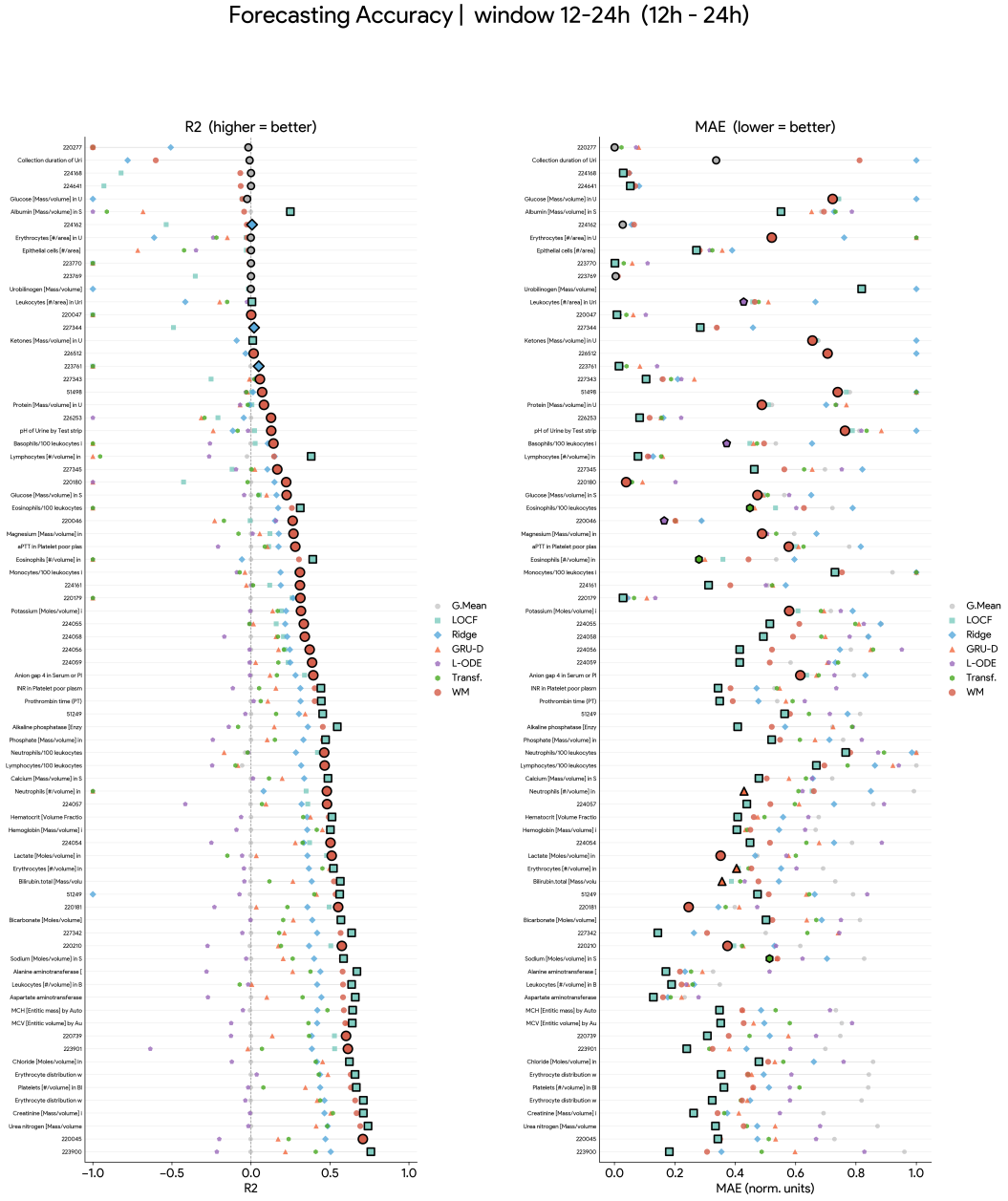

Figure 22: **Regression decoding performance at long-horizon rollout (12–24h).** Per-feature evaluation of forecasting accuracy for the 12–24 hour prediction window, with features sorted by world model (WM) R^2 in descending order. Left: coefficient of determination (R^2); right: mean absolute error (MAE).

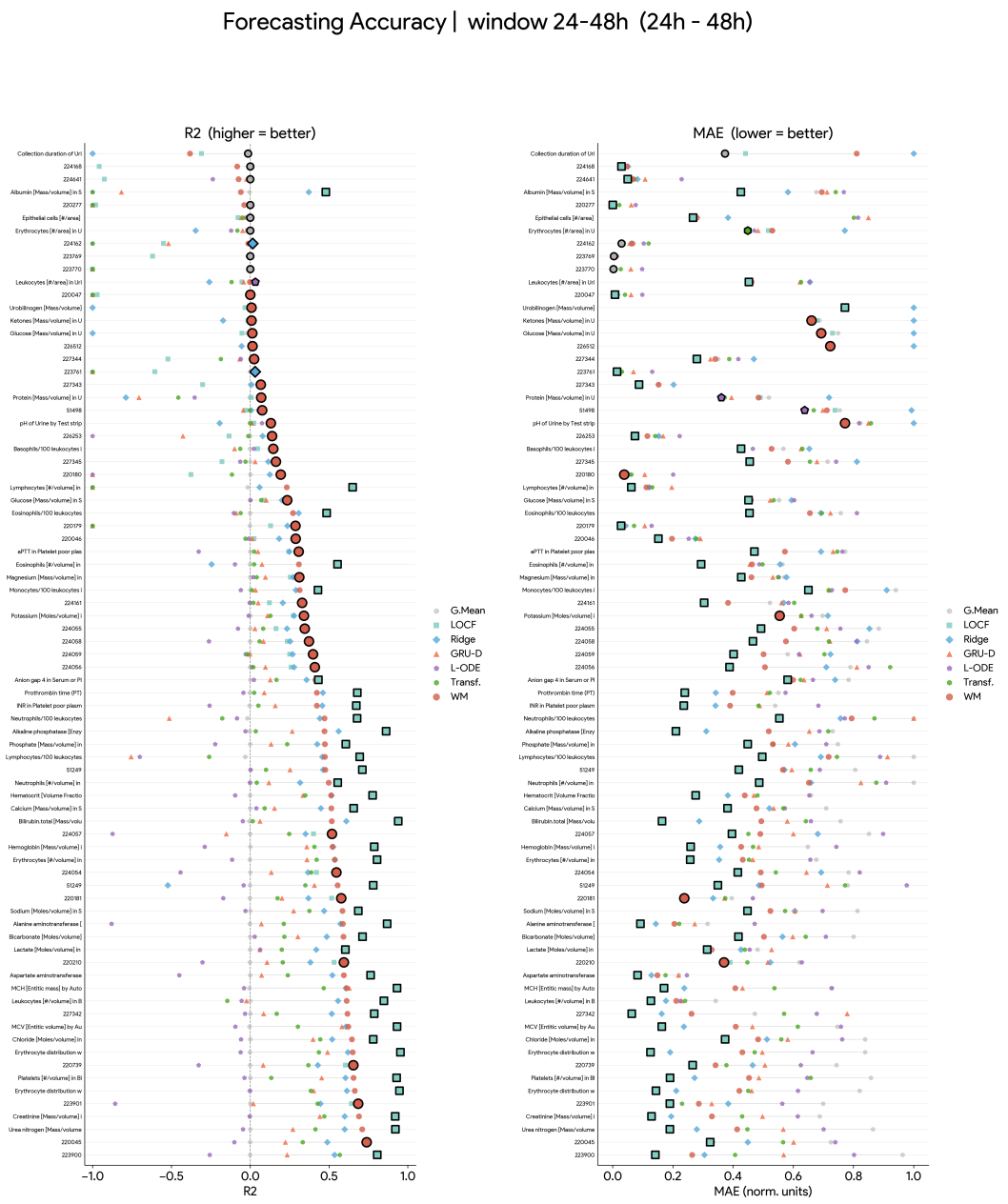

Figure 23: **Regression decoding performance at long-horizon rollout (24–48h).** Per-feature evaluation of forecasting accuracy for the 24–48 hour prediction window, with features sorted by world model (WM) R^2 in descending order. Left: coefficient of determination (R^2); right: mean absolute error (MAE).