# DECAF: DECOMPOSED SUBSPACES FOR LONGITUDINAL TREATMENT EFFECT ESTIMATION

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Estimating how a patient's outcomes might change under different treatment plans over time is a central need in personalized medicine. Longitudinal observational data, especially from electronic health records (EHR), can help address this by capturing how clinical decisions and patient states evolve across repeated visits. In real-world settings, however, treatment practices and data collection processes often vary across environments (e.g., across hospitals, cohorts or over time). As a result, signals that reflect the patient's underlying clinical state and treatment response, i.e., the factors through which interventions causally influence outcomes, can be entangled with environment-specific artifacts: systematic, setting-induced patterns (e.g., differences in treatment protocols and documentation practices), making multi-step counterfactual forecasting unreliable. To address this challenge, we propose DECAF, a framework that learns a structured decomposed representation by separating a patient's longitudinal history into two latent subspaces: an outcome-relevant subspace that encodes stable, intervention-relevant factors, and an environment/nuisance subspace that captures variation associated with treatment practice, cohort, and measurement processes. We then combine this representation with a trajectory generator to produce counterfactual trajectories under alternative treatment sequences, enabling coherent and stable forecasts over time.

## 1 INTRODUCTION

Estimating *individualized treatment effects* (ITE) from longitudinal observational data is central to sequential decision-making in healthcare, where treatments are assigned repeatedly and patient states evolve over time (Robins et al., 2000; Hernán & Robins, 2010; Lim, 2018; Bica et al., 2020b). A practical goal is to predict how a patient's future outcome trajectory would change under different treatment plans, and to quantify these effects across multiple forecasting horizons.

A major challenge in this setting is *policy and environment shift*. In real clinical data, treatment policies, measurement processes, and patient populations can vary across hospitals, cohorts, and time, causing models trained in one environment to fail under another (Peters et al., 2016; Arjovsky et al., 2019; Schölkopf et al., 2021; Liu et al., 2023). A model may therefore learn correlations that reflect local practice rather than stable treatment response. For example, a treatment may appear predictive of worse outcomes in one hospital because it is used mostly for severe patients, not because the treatment itself causes deterioration. Such environment-specific shortcuts can make counterfactual predictions unreliable under a different policy or cohort (Liu et al., 2020; 2023). This motivates representations that preserve information relevant to treatment response while reducing dependence on environment-specific variation, in line with causal representation learning and invariance-based generalization (Peters et al., 2016; Arjovsky et al., 2019; Schölkopf et al., 2021).

Existing deep longitudinal ITE methods usually encode patient history into a single latent representation and use it to predict future outcomes under hypothetical treatments (Lim, 2018; Bica et al., 2020b;a; Melnychuk et al., 2022). Recent generative and contrastive approaches further improve counterfactual coherence by sampling or aligning counterfactuals in latent space (Louizos et al., 2017; Yoon et al., 2018; Zhang et al., 2025; Wu et al., 2025a;b). However, a single latent representation can mix treatment-relevant patient state with environment-related artifacts, such as site-specific measurement patterns, cohort differences, or local documentation practices. When counterfactual rollouts are generated from this entangled representation, the model may perturb both the treatment-relevant state



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

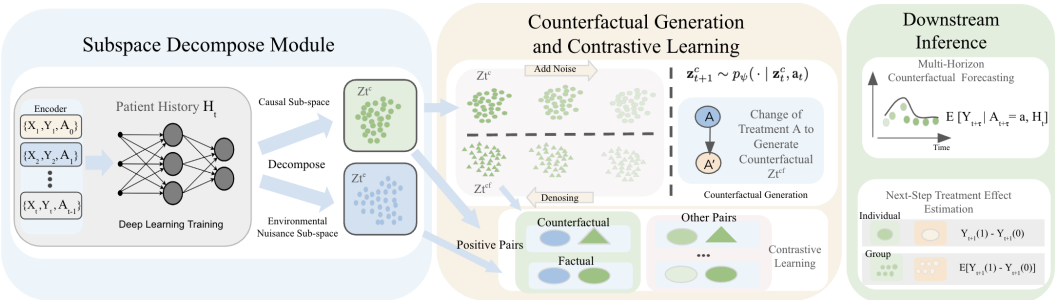


Figure 1: **Overview of DECAF.** The encoder maps history $\mathbb{H}_t$ to $(\mathbf{z}_t^c, \mathbf{z}_t^e)$; outcome heads use $\mathbf{z}_t^c$. A treatment-conditional diffusion model translates $\mathbf{z}_t^c$ under an intervened $\mathrm{a}_t'$ while $\mathbf{z}_t^e$ is held fixed, producing factual/counterfactual pairs that support multi-horizon forecasting and ITE.

and the nuisance variation, which can amplify errors over multiple forecasting steps (Rasul et al., 2021; Song et al., 2020).

To address this issue, we introduce **DECAF**, a longitudinal counterfactual forecasting framework that decomposes each patient's history into two latent components. Here component $z_t^c$ is encouraged to retain information needed for causal treatment-effect prediction and treatment-conditioned rollout, while the environment component $z_t^e$ is encouraged to capture variation associated with local practice patterns, measurement processes, and cohort differences. At inference time, DECAF evolves future $z_t^c$ under alternative treatment plans while keeping $z_t^e$ fixed, allowing predictions to vary with the intervened treatment sequence without directly modifying environment-related variation. This design helps reduce reliance on site-, cohort-, or policy-specific shortcuts under policy or environment shift (Bousmalis et al., 2016; Locatello et al., 2020; Schölkopf et al., 2021).

The contributions of this work are threefold. **a) Decomposed longitudinal representation.** We propose a representation learning objective that separates outcome-relevant signals from environment-related variation for robust treatment effect estimation under shift. **b) Latent counterfactual rollout.** We introduce a treatment-conditional diffusion model that performs multi-step counterfactual forecasting in the outcome-relevant latent space. **c) Evaluation under policy and environment shift.** We evaluate DECAF on the Tumor Growth simulator (Lim, 2018; Melnychuk et al., 2022), semi-synthetic MIMIC-III (Liu et al., 2020; Melnychuk et al., 2022), and a real MIMIC-IV mechanical ventilation cohort (Johnson et al., 2023; Moukheiber et al., 2024a).

## 2 RELATED WORK

**Longitudinal treatment effect estimation.** Longitudinal causal inference studies treatment effects under sequential decisions and time-varying confounding. Classical approaches, such as marginal structural models and the sequential g-formula, rely on reweighting or outcome modeling under sequential ignorability (Robins et al., 2000; Hernán & Robins, 2010). Recent deep methods learn history representations for counterfactual prediction, including RMSNs (Lim, 2018), CRN (Bica et al., 2020b), Time Series Deconfounder (Bica et al., 2020a), SyncTwin (Qian et al., 2021), G-Net (Li et al., 2020), Causal Transformer (Melnychuk et al., 2022), and continuous-time extensions such as TE-CDE (Seedat et al., 2022). Structured representation methods such as DCRN (Berrevoets et al., 2021) and DR-CRN (Chu et al., 2022) are closest to DECAF. Unlike these methods, DECAF ties the decomposition to environment and policy shift: $z_t^c$ supports potential-outcome prediction and latent rollout, while $z_t^e$ absorbs environment-, cohort-, and measurement-related variation.

**Causal and invariant representation learning.** Representation learning for causal inference aims to reduce treated–control mismatch or improve robustness across environments (Johansson et al., 2016; Shalit et al., 2017; Yao et al., 2018; Schölkopf et al., 2021). Related ideas include invariant prediction (Peters et al., 2016), invariant risk minimization (Arjovsky et al., 2019), group distributional robustness (Sagawa et al., 2019), domain-adversarial learning (Ganin et al., 2016), and domain separation networks (Bousmalis et al., 2016). Meaningful factorization often requires supervision or structural inductive bias (Locatello et al., 2020).

For longitudinal treatment-effect estimation, DECAF assigns its two latent components explicit prediction and environment-separation roles.

**Generative counterfactual modeling.** Generative models have been used to synthesize missing potential outcomes and counterfactual samples, including latent-variable models, adversarial models, and normalizing flows (Louizos et al., 2017; Yoon et al., 2018; Javaloy et al., 2023; Dombrowski et al., 2023; Wu et al., 2025a). Recent work extends this direction to robust ITE estimation and temporal counterfactuals through flow-based, contrastive, and time-varying treatment models (Zhang et al., 2025; Wu et al., 2024; 2025b; Wang et al., 2023). Diffusion models have also been effective for probabilistic time-series forecasting (Rasul et al., 2021; Ho et al., 2020; Song et al., 2020), with recent work targeting causal counterfactual outcome distributions in longitudinal data (Alinezhad et al., 2026). DECAF differs by applying treatment-conditional diffusion only in the outcome-relevant latent subspace, rather than generating counterfactual outcomes or full entangled trajectories directly. A full related work is provided in Appendix C.

## 3 PRELIMINARY

**Study setting and notation.** We consider a longitudinal observational dataset with $N$ units, such as patients, observed over time. For unit $n$, let $T^{(n)}$ denote its trajectory length. At each time step $t \in \{1, \dots, T^{(n)}\}$, we observe time-varying covariates $\mathbf{x}_t \in \mathbb{R}^d$, a treatment $\mathrm{a}_t \in \mathcal{A}$, and an outcome $\mathrm{y}_t \in \mathbb{R}$. For binary-treatment settings, $\mathcal{A} = \{0, 1\}$; for settings with multiple binary treatments, we encode the joint treatment as a categorical treatment in $\mathcal{A}$. We define the observed history available before assigning treatment $\mathrm{a}_t$ as:

$$\mathbb{H}_t := \{\mathbf{X}_t, \boldsymbol{a}_{t-1}, \boldsymbol{y}_t\},$$

where $\mathbf{X}_t = (\mathbf{x}_1, \dots, \mathbf{x}_t) \in \mathbb{R}^{t \times d}$ is the covariate history, $\boldsymbol{a}_{t-1} = (\mathrm{a}_1, \dots, \mathrm{a}_{t-1}) \in \mathcal{A}^{t-1}$ is the past treatment history, and $\boldsymbol{y}_t = (\mathrm{y}_1, \dots, \mathrm{y}_t) \in \mathbb{R}^t$ is the observed outcome history. After treatment $\mathrm{a}_t$ is assigned, the next outcome $\mathrm{y}_{t+1}$ is observed.

**Causal estimands and learning goals.** Let $\mathrm{y}_{t+1}^{\boldsymbol{a}_t}$ denote the potential outcome at time $t+1$ under the treatment sequence $\boldsymbol{a}_t = (a_1, \dots, a_t)$ (Rubin, 1974; Robins, 1986; Hernán & Robins, 2010). Under the standard longitudinal identification assumptions, including consistency, sequential ignorability, positivity, and no interference (see full definitions in Appendix H.1), the conditional mean potential outcome is identifiable from observational data as:

$$\mathbb{E}[\mathrm{y}_{t+1}^{\boldsymbol{a}_t} \mid \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbb{H}_t], \qquad a \in \mathcal{A},$$

where past treatments are already included in $\mathbb{H}_t$ and $\boldsymbol{a}_t = (\boldsymbol{a}_{t-1}, a)$. For clarity, we state the step-wise individualized treatment effect for binary treatment:

$$\tau_t(\mathbb{H}_t) := \mathbb{E}\left[\mathrm{y}_{t+1}^1 - \mathrm{y}_{t+1}^0 \mid \mathbb{H}_t\right],$$

where $\mathrm{y}_{t+1}^a$ is the potential outcome if treatment $a \in \{0, 1\}$ were assigned at time $t$, holding the observed history fixed. For finite multi-treatment settings, we evaluate potential outcomes and counterfactual rollouts for each $a \in \mathcal{A}$ or for specified future treatment sequences.

The dataset is written as $\mathcal{D} = \{(\mathbf{X}_{T^{(n)}}^{(n)}, \boldsymbol{a}_{T^{(n)}}^{(n)}, \boldsymbol{y}_{T^{(n)}}^{(n)}, \mathrm{e}^{(n)})\}_{n=1}^N$, where $\mathrm{e}^{(n)} \in \mathcal{E}$ denotes the environment of trajectory $n$, such as the hospital, cohort, time period, or treatment policy. When available, $\mathrm{e}^{(n)}$ is used during training to encourage robustness to environment shift. We evaluate two tasks. **T1: Next-step treatment-effect estimation.** Given $\mathbb{H}_t$, estimate $(\widehat{\mathrm{y}}_{t+1}^1, \widehat{\mathrm{y}}_{t+1}^0)$ and $\widehat{\tau}_t = \widehat{\mathrm{y}}_{t+1}^1 - \widehat{\mathrm{y}}_{t+1}^0$ for binary settings. **T2: Multi-horizon counterfactual forecasting.** Given $\mathbb{H}_t$ and a future treatment sequence, predict the potential outcome trajectory over multiple horizons.

**Motivation and challenges.** Longitudinal counterfactual prediction requires accurate potential-outcome estimation and multi-step rollout under alternative treatment plans (Robins et al., 2000; Lim, 2018; Bica et al., 2020b; Melnychuk et al., 2022). A single history representation may entangle patient state with measurement, cohort, or policy information, encouraging unstable shortcuts under shift (Peters et al., 2016; Arjovsky et al., 2019; Geirhos et al., 2020; Schölkopf et al., 2021). DECAF adapts domain separation by learning outcome-relevant $\mathbf{z}_t^c$ and environment-related $\mathbf{z}_t^e$, then generating treatment-conditioned futures through $\mathbf{z}_t^c$ (Bousmalis et al., 2016; Locatello et al., 2020).

# 4 METHODOLOGY

**Overview.** DECAF learns a representation for outcome prediction and treatment-conditioned forecasting. Stage 1 encodes the pre-treatment history $\mathbb{H}_t = (\mathbf{X}_t, \boldsymbol{a}_{t-1}, \boldsymbol{y}_t)$ as $(\mathbf{z}_t^c, \mathbf{z}_t^e) = f_{\theta_f}(\mathbb{H}_t)$, with outcome heads using $\mathbf{z}_t^c$ and an auxiliary branch $\mathbf{z}_t^e$ retaining environment information. Stage 2 freezes the encoder and heads and learns transitions of $\mathbf{z}_t^c$. Environment labels are needed only during training. Algorithms 1–3 give the full procedure.

## 4.1 STAGE 1: DECOMPOSED REPRESENTATION LEARNING

The branches have dimensions $d_c$ and $d_e$. Each treatment $a \in \mathcal{A}$ has an outcome head $g_a$. The finite alphabet can include joint treatments: chemotherapy and radiotherapy, for example, give four binary-treatment combinations. Below, $\widehat{\mathbb{E}}$ denotes the minibatch mean.

**Prediction and environment separation.** We optimize the outcome and environment objectives

$$\mathcal{L}_y = \widehat{\mathbb{E}}[(\mathrm{y}_{t+1} - g_{\mathrm{a}_t}(\mathbf{z}_t^c))^2],$$
$$\mathcal{L}_{\mathrm{env}}^b = \widehat{\mathbb{E}}\, \ell_{\mathrm{CE}}(\mathrm{e}, h_{\mathrm{env}}^b(\mathbf{z}_t^b)), \qquad b \in \{c, e\}.$$

Each example supervises its observed treatment head. Both classifiers minimize their losses, while the encoder reverses the gradient from $\mathcal{L}_{\mathrm{env}}^c$ and follows that from $\mathcal{L}_{\mathrm{env}}^e$ (Ganin et al., 2016; Bousmalis et al., 2016). This discourages environment prediction from $\mathbf{z}_t^c$ while retaining it in $\mathbf{z}_t^e$.

**Preserving outcome information.** Environment information can also predict outcomes: patient severity, for example, may distinguish ICU populations as well as predict their outcomes. We therefore compare treatment-conditioned probes using $\mathbf{z}_t^c$ alone and using $(\mathbf{z}_t^c, \mathbb{H}_t)$:

$$\ell_c = \widehat{\mathbb{E}}[(\mathrm{y}_{t+1} - h_{\mathrm{suf},\mathrm{a}_t}^c(\mathbf{z}_t^c))^2],$$
$$\ell_{\mathrm{aug}} = \widehat{\mathbb{E}}[(\mathrm{y}_{t+1} - h_{\mathrm{suf},\mathrm{a}_t}^{\mathrm{aug}}(\mathbf{z}_t^c, \mathbb{H}_t))^2],$$
$$\mathcal{L}_{\mathrm{suf}} = [\ell_c - \mathrm{sg}(\ell_{\mathrm{aug}}) - \delta]_+, \qquad [x]_+ = \max(0, x), \quad \delta \geq 0.$$

The probes first minimize $\ell_c + \ell_{\mathrm{aug}}$ with detached encoder outputs. We then fix their parameters and recompute the losses for the encoder update, retaining input gradients through the latent-only probe. Stopping gradients through the augmented loss makes it a reference that the encoder cannot directly worsen to reduce the penalty.

**Alternating optimization.** With classifiers and probes fixed, the encoder minimizes

$$\mathcal{J}_f = \mathcal{L}_y + \lambda_{\mathrm{env}}(\mathcal{L}_{\mathrm{env}}^e - \mathcal{L}_{\mathrm{env}}^c) + \lambda_{\mathrm{suf}}\mathcal{L}_{\mathrm{suf}}.$$

Outcome heads minimize $\mathcal{L}_y$; environment classifiers minimize $\mathcal{L}_{\mathrm{env}}^c + \mathcal{L}_{\mathrm{env}}^e$ on detached features. The weights balance environment separation against retention of predictive information.

## 4.2 STAGE 2: TREATMENT-CONDITIONAL LATENT TRANSITIONS

With the encoder and outcome heads frozen, adjacent factual histories provide transition examples $(\mathbf{z}_t^c, \mathrm{a}_t, \mathbf{z}_{t+1}^c)$. A conditional denoiser and DDIM sampler define $p_\psi(\cdot \mid \mathbf{z}_t^c, \mathrm{a}_t)$.

**Factual diffusion training.** Let $\alpha_k = 1 - \beta_k$ and $\bar{\alpha}_k = \prod_{j=1}^k \alpha_j$. Independently draw $k \sim \mathrm{Unif}\{1, \ldots, K\}$ and $\epsilon \sim \mathcal{N}(0, \mathbf{I})$, and use the DDPM objective (Ho et al., 2020):

$$\mathbf{z}_{t+1,k}^c = \sqrt{\bar{\alpha}_k}\mathbf{z}_{t+1}^c + \sqrt{1 - \bar{\alpha}_k}\epsilon,$$
$$\mathcal{L}_{\mathrm{diff}} = \mathbb{E}_{(\mathbb{H}_t, \mathrm{a}_t, \mathbb{H}_{t+1}), k, \epsilon}\left[\left\|\epsilon - \epsilon_\psi(\mathbf{z}_{t+1,k}^c, k \mid \mathbf{z}_t^c, \mathrm{a}_t)\right\|_2^2\right].$$

Factual minibatches approximate the transition expectation. The denoiser and reverse sampler jointly determine the generated distribution (Song et al., 2021); Appendix B details sampling.

**Same-history pair regularization.** For a given history, alternative treatments may produce different responses, as with chemotherapy alone or combined with radiotherapy. We use projections to encourage shared patient information across these transitions. For each example, draw $a'_t \sim \mathrm{Unif}(\mathcal{A} \setminus \{\mathrm{a}_t\})$ and form

$$\widetilde{\mathbf{z}}^c_{t+1} = G_\psi(\mathbf{z}^c_t, a'_t; \xi), \qquad \xi \sim \mathcal{N}(0, \mathbf{I}),$$
$$u_i = h([\mathbf{z}^{c,(i)}_{t+1}; \mathbf{z}^{e,(i)}_t]), \qquad v_i = q([\widetilde{\mathbf{z}}^{c,(i)}_{t+1}; \mathbf{z}^{e,(i)}_t]).$$

The differentiable DDIM sampler $G_\psi$ uses independent base-noise draws $\xi$. Each minibatch contains distinct patients. With cosine similarity, we minimize (Oord et al., 2018; Chen et al., 2020)

$$\mathcal{L}_{\mathrm{con}} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\mathrm{sim}(u_i, v_i)/\kappa_{\mathrm{con}})}{\sum_{j=1}^{B} \exp(\mathrm{sim}(u_i, v_j)/\kappa_{\mathrm{con}})}.$$

Stage 2 updates $\psi, h, q$ with $\mathcal{L}_{\mathrm{diff}} + \lambda_{\mathrm{con}}\mathcal{L}_{\mathrm{con}}$, retaining gradients through $G_\psi$. Alignment acts on projections, allowing treatment-specific latents and outcomes. Because shared $\mathbf{z}^e_t$ can facilitate alignment, low pair loss alone does not establish transition accuracy or treatment-effect preservation.

**Multi-step prediction.** For a fixed plan $\bar{a} = (a_t, \ldots, a_{t+\tau-1})$, initialize $M$ paths at $\mathbf{z}_t^{c,(m)} = f^c_{\theta_f}(\mathbb{H}_t)$. At each time $s = t, \ldots, t+\tau-1$, predict the next outcome and sample a transition only if another horizon is required:

$$\widehat{\mu}^{\bar{a}}_{s+1}(\mathbb{H}_t) = \frac{1}{M}\sum_{m=1}^{M} g_{a_s}(\mathbf{z}^{c,(m)}_s),$$
$$\mathbf{z}^{c,(m)}_{s+1} \sim p_\psi(\cdot \mid \mathbf{z}^{c,(m)}_s, a_s), \qquad s < t+\tau-1.$$

Sampling noise is independent across paths and transitions. Only $\mathbf{z}^c$ is propagated; one-step prediction is $g_{a_t}(\mathbf{z}^c_t)$. Differences between plan means estimate conditional mean treatment effects under the conditions in Section 5. Head outputs represent conditional means, not full outcome draws.

## 5 Theory for Representation Sufficiency

We state population conditions for one-step identification, multi-step rollout, and transfer. They clarify what the representation must retain; the training objectives encourage, but do not guarantee, these properties. Proofs are in Appendix H.

**Assumption 1** (Conditional-mean sufficiency)**.** Let $(\mathbf{z}^c_t, \mathbf{z}^e_t) = f_t(\mathbb{H}_t)$ be measurable. For each $a \in \mathcal{A}$, suppose $\mathrm{y}^a_{t+1}$ has finite second moment and

$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbb{H}_t] = m_{t,a}(\mathbf{z}^c_t) \quad \text{a.s.}$$

for a measurable function $m_{t,a}$.

This condition concerns treatment-specific means, allowing dependence between the branches without requiring recovery of a physical state. Given $\mathbf{z}^c_t$, $\mathbf{z}^e_t$ adds no information about these means.

**Theorem 1** (Identification through $\mathbf{z}^c_t$)**.** *Assume consistency,* $\mathrm{y}^a_{t+1} \perp\!\!\!\perp \mathrm{a}_t \mid \mathbb{H}_t$*, and* $P(\mathrm{a}_t = a \mid \mathbb{H}_t) > 0$ *for each evaluated treatment on the history support of interest. Under Assumption 1,*

$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbb{H}_t] = \mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbf{z}^c_t] = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}^c_t] \quad \textit{a.s.}$$

Thus outcome heads can use $\mathbf{z}^c_t$ in place of history when mean sufficiency holds. Identification still requires adequate history-based confounding adjustment; environment regularization does not remove residual hidden confounding.

**Connection to the probes.** For a fixed time and representation, let $\mu_H = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathbb{H}_t, \mathrm{a}_t]$ and $\mu_Z = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathbf{z}^c_t, \mathrm{a}_t]$. The optimal squared risks over all measurable predictors satisfy

$$R^\star_Z - R^\star_H = \mathbb{E}[(\mu_H - \mu_Z)^2]. \tag{1}$$

Under positivity, zero gap gives treatment-specific conditional-mean equality on supported histories. This motivates the probe comparison, whose accuracy depends on capacity, optimization, and slack. A small average gap can still conceal errors on rare treatment histories.

**Multi-step rollout.** Accurate next-step means need not determine future dynamics. For example, histories supporting similar next-hour oxygenation predictions may differ in prior ventilation and subsequent trajectories. Define $\mathbf{z}_{s+1}^{c,a} = f_{s+1}^{c}(\mathbb{H}_{s+1}^{a})$, where $\mathbb{H}_{s+1}^{a}$ is the next history under treatment $a$. In addition to mean sufficiency at future decision times, require

$$\mathcal{L}(\mathbf{z}_{s+1}^{c,a} \mid \mathbb{H}_s) = K_{s,a}(\cdot \mid \mathbf{z}_s^c). \tag{2}$$

For a fixed plan $\bar{a} = (a_t, \ldots, a_{t+\tau-1})$, assume consistency, sequential exchangeability for the future covariate and outcome process, positivity, and integrable target outcomes, with all conditions holding on histories reached by the plan. Then $K_{s,a}(\cdot \mid \mathbf{z}_s^c) = \mathcal{L}(\mathbf{z}_{s+1}^c \mid \mathrm{a}_s = a, \mathbf{z}_s^c)$, and

$$\mathbb{E}[\mathrm{y}_{t+\tau}^{\bar{a}} \mid \mathbb{H}_t] = \int m_{t+\tau-1,a_{t+\tau-1}}(z_{t+\tau-1}) \prod_{s=t}^{t+\tau-2} K_{s,a_s}(dz_{s+1} \mid z_s), \tag{3}$$

where $z_t = \mathbf{z}_t^c$; for $\tau = 1$, this reduces to $m_{t,a_t}(z_t)$. The heads and sampler approximate these functions and kernels; their time-shared implementation requires temporal information in the encoded state. Transition sufficiency is additional to the one-step probe condition. Monte Carlo averaging reduces sampling error for the fitted model, while representation and transition errors remain.

**Transfer across environments.** A shared outcome predictor requires

$$\mathbb{E}[\mathrm{y}_{t+1}^{a} \mid \mathbf{z}_t^c = z, \mathrm{e} = e] = m_{t,a}(z)$$

on the relevant support. Multi-step transfer also requires shared conditional transition kernels and support overlap for target histories and treatment plans. ICU units can differ in patient mix even when conditional response mechanisms agree. Thus marginal invariance, encouraged by the adversary, neither establishes these mechanisms nor guarantees sufficiency; removing environment information can also remove predictive patient characteristics.

**A stylized subspace result.** Under the Gaussian-state, noise, rank, and rich-environment assumptions in Appendix H.3, the linear model $\mathbf{h}_t = \mathbf{A}\mathbf{s}_t + \mathbf{B}\mathbf{u}_t$ illustrates the role of each objective. Marginal invariance of $\mathbf{z}_t^c = \mathbf{W}^\top \mathbf{h}_t$ implies $\mathbf{W}^\top \mathbf{B} = 0$, while exact outcome sufficiency preserves the treatment-specific linear outcome functionals. Full-state recovery additionally requires these functionals to span the state. This illustrates compatibility without verifying the assumptions in the benchmarks or guaranteeing nonlinear recovery.

## 6 Experiments

We evaluate retention of treatment-response information, multi-step prediction with latent diffusion, and transfer to held-out environments. Tumor Growth and semi-synthetic MIMIC-III provide oracle counterfactuals; real MIMIC-IV assesses factual forecasting.

### 6.1 Evaluation setup

Baselines cover weighting, outcome regression, and longitudinal sequence modeling: MSM-IPW (Robins et al., 2000), TARNet (Shalit et al., 2017), RMSN (Lim, 2018), CRN (Bica et al., 2020b), DR-CRN (Chu et al., 2022), G-Net (Li et al., 2020), Causal Transformer (Melnychuk et al., 2022), G-Transformer (Xiong et al., 2024), and CTD-NKO (Wang et al., 2025). We also include the Causal Diffusion Model (CDM; Alinezhad et al., 2026), which generates counterfactual outcome distributions. Real MIMIC-IV additionally includes TE-CDE (Seedat et al., 2022) and Persistence, which holds the last outcome fixed. Appendix F details the baselines evaluated in each table.

We report endpoint normalized RMSE for forecasting, $\sqrt{\mathrm{PEHE}}$ for individual effects, and ATEErr for average effects; lower is better (Appendix G). Experiments use seeds 10, 101, and 1010. DeCaf averages 100 latent paths with 20 DDIM updates per transition; at $\tau = 1$, only the Stage 1 encoder and outcome heads are used (Appendix A.4).

Following Appendix D.1, target environments are excluded from training, validation, model selection, and normalization. Tumor Growth models train jointly on $\gamma \in \{0.5, 1.0, 1.5\}$. Held-out patients at $\gamma = 1$ represent ID; $\gamma = 2, 3, 4$ define OOD-1, OOD-2, and OOD-3. Models are evaluated without adaptation. Querying counterfactual treatments for a source-environment history does not itself constitute environment OOD.

### 6.2 TUMOR GROWTH: FORECASTING UNDER TREATMENT PLANS

The simulator evaluates chemotherapy and radiotherapy plans (Geng et al., 2017; Lim, 2018; Bica et al., 2020b; Melnychuk et al., 2022). The policy coefficient $\gamma$ controls treatment assignment from recent tumor measurements, while response mechanisms remain fixed across environments. Radiotherapy acts immediately and chemotherapy has a delayed effect. One-step prediction uses the representation and outcome heads; later horizons also require latent transitions.

Table 1: **Tumor Growth forecasting.** $\mathrm{nRMSE}_{cf}(\tau) \downarrow$ on held-out patients under the representative ID policy $\gamma = 1$ and OOD-2 policy $\gamma = 3$. Entries are mean $\pm$ sample standard deviation over three seeds. Bold marks the lowest mean in each setting.

| $\gamma$ | Method | $\tau = 1$ | $\tau = 2$ | $\tau = 3$ | $\tau = 4$ | $\tau = 5$ |
|---|---|---|---|---|---|---|
| 1 | MSM | 2.274±0.376 | 2.787±0.451 | 3.087±0.499 | 3.222±0.519 | 3.216±0.501 |
| | RMSN | 1.446±0.225 | 1.524±0.253 | 1.583±0.259 | 1.612±0.243 | 1.611±0.197 |
| | CRN | 1.300±0.127 | 1.371±0.173 | 1.431±0.196 | 1.477±0.206 | 1.520±0.205 |
| | DR-CRN | 1.693±0.121 | 1.732±0.158 | 1.761±0.179 | 1.766±0.194 | 1.746±0.177 |
| | G-Net | 1.655±0.197 | 1.949±0.307 | 2.162±0.387 | 2.353±0.451 | 2.551±0.510 |
| | CT | 1.593±0.094 | 1.709±0.096 | 1.790±0.113 | 1.828±0.125 | 1.805±0.129 |
| | G-Transformer | 2.235±0.263 | 2.699±0.349 | 3.173±0.351 | 3.362±0.489 | 3.437±0.380 |
| | CTD-NKO | 1.308±0.042 | 1.358±0.073 | 1.416±0.109 | 1.475±0.132 | 1.538±0.130 |
| | CDM | 1.203±0.013 | 1.237±0.046 | 1.412±0.204 | 1.457±0.176 | 1.578±0.175 |
| | **DECAF** | **0.930±0.114** | **1.068±0.159** | **1.231±0.376** | **1.393±0.537** | **1.515±0.624** |
| 3 | MSM | 3.149±0.491 | 3.543±0.583 | 3.780±0.630 | 3.892±0.652 | 3.857±0.627 |
| | RMSN | 1.675±0.332 | 1.820±0.403 | 2.007±0.549 | 2.121±0.597 | 2.129±0.561 |
| | CRN | 1.920±0.427 | 2.237±0.536 | 2.393±0.569 | 2.447±0.568 | 2.405±0.525 |
| | DR-CRN | 1.680±0.223 | 1.958±0.369 | 2.200±0.420 | 2.383±0.424 | 2.514±0.385 |
| | G-Net | 1.940±0.120 | 2.459±0.161 | 2.870±0.314 | 3.219±0.566 | 3.510±0.892 |
| | CT | 2.194±0.203 | 2.536±0.320 | 2.691±0.357 | 2.745±0.418 | 2.654±0.434 |
| | G-Transformer | 3.428±0.574 | 4.263±0.633 | 4.828±0.750 | 5.063±0.691 | 4.970±0.641 |
| | CTD-NKO | 1.561±0.136 | 1.738±0.198 | 1.885±0.242 | 1.995±0.264 | 2.075±0.264 |
| | CDM | 1.521±0.858 | 1.631±0.012 | 1.929±0.258 | 2.942±0.271 | 2.578±0.824 |
| | **DECAF** | **1.431±0.297** | **1.275±0.166** | **1.439±0.263** | **1.770±0.359** | **2.067±0.468** |

DECAF has the lowest mean error at every horizon for $\gamma = 1$ and $\gamma = 3$ in Table 1. The advantage is small at $\tau = 5$, where the differences from the best baseline are $0.005$ (ID) and $0.008$ (OOD-2), respectively. Across the three held-out policies in Appendix D.3, it ranks first in 13 of 15 OOD settings (18 of 20 when including the five representative ID settings); CTD-NKO has lower means at $(\gamma, \tau) = (2, 5)$ and $(4, 1)$. At $(4, 5)$, DECAF achieves $2.277$, compared with $2.707$ for the best baseline. The gains therefore vary across settings, with a long-horizon advantage in the strongest confounding setting considered ($15.9\%$ at $\gamma = 4$, $\tau = 5$). Near-ties at the other reported endpoints should be read alongside the seed standard deviations, rather than as uniform superiority.

### 6.3 REPRESENTATION AND TRANSITION ANALYSIS

Environment-related features such as patient severity may also predict treatment response. Stage 1 balances environment separation with retention of outcome information. Figure 2 visualizes the learned branches; the following comparisons assess their predictive use.

Table 2 compares DDPM (Ho et al., 2020), CVAE (Kingma & Welling, 2013; Sohn et al., 2015), normalizing flow (Rezende & Mohamed, 2015), and Gaussian transitions. The comparison starts at $\tau = 2$; one-step prediction uses Stage 1 alone.

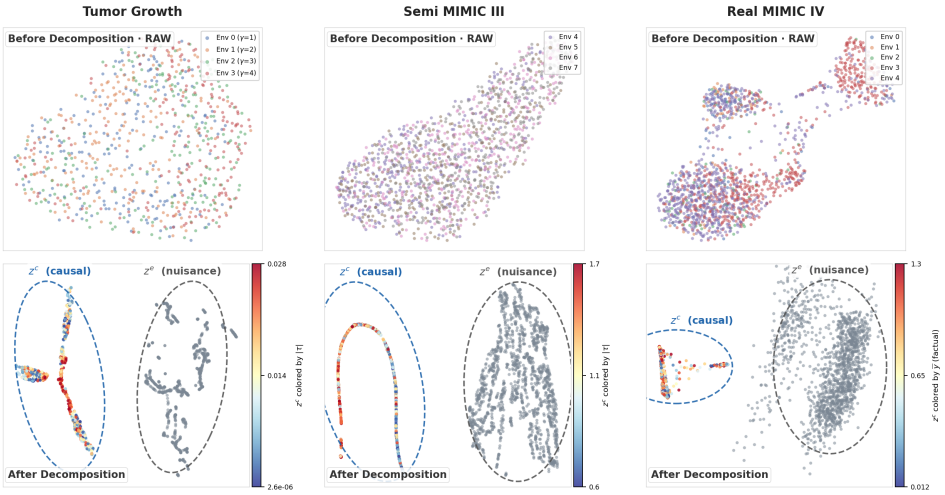


Figure 2: **Representations across datasets.** Top: raw latents colored by environment. Bottom: the learned branches $\mathbf{z}^c$ and $\mathbf{z}^e$. Colors in $\mathbf{z}^c$ indicate oracle treatment-effect magnitude for Tumor Growth and semi-synthetic MIMIC-III, and observed outcomes for real MIMIC-IV.

Table 2: **Transition family and latent representation on Tumor Growth.** Mean $\mathrm{nRMSE}_{cf}(\tau) \downarrow$ over the ID policy $\gamma = 1$, OOD policies $\gamma \in \{2, 3, 4\}$, and three seeds. Within each latent group, columns report $\tau = 2, 3, 4, 5$. Bold marks the lowest mean at each horizon.

| | Outcome-relevant $z^c$ | | | | Unsplit $z$ | | | | Environment $z^e$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Transition | 2 | 3 | 4 | 5 | 2 | 3 | 4 | 5 | 2 | 3 | 4 | 5 |
| DDPM | **1.394** | **1.525** | **1.714** | **1.880** | 1.483 | 1.618 | 1.912 | 2.175 | 28.151 | 26.470 | 25.963 | 25.773 |
| CVAE | 1.504 | 1.765 | 1.985 | 2.200 | 1.981 | 2.471 | 2.271 | 2.654 | 29.452 | 27.644 | 26.452 | 21.389 |
| Flow | 1.787 | 2.190 | 2.534 | 2.839 | 1.924 | 2.654 | 2.812 | 2.971 | 34.584 | 34.399 | 34.188 | 33.924 |
| Gaussian | 2.185 | 2.639 | 3.045 | 3.439 | 2.427 | 2.917 | 3.467 | 3.671 | 31.589 | 28.781 | 26.302 | 24.490 |

**Transition family.** Holding $\mathbf{z}^c$ fixed, DDPM improves on CVAE, the strongest alternative in this subspace, by $7.3\%$ at $\tau = 2$ ($1.394$ versus $1.504$) and $14.5\%$ at $\tau = 5$ ($1.880$ versus $2.200$). The advantage thus widens over the evaluated horizons.

**Representation.** Holding DDPM fixed, replacing the unsplit state with $\mathbf{z}^c$ lowers error by $6.0\%$ at $\tau = 2$ and $13.6\%$ at $\tau = 5$. The absolute error increase from $\tau = 2$ to $5$ is $0.486$ in $\mathbf{z}^c$, compared with $0.692$ in the unsplit state. Poor environment-only prediction does not show that $\mathbf{z}^e$ contains no outcome information. These mixed ID/OOD means do not isolate policy-shift effects.

**Training objectives.** Loss ablations in Appendix E assess both stages. At $\tau = 1$, removing environment separation or sufficiency raises the four-policy mean error from $1.371$ to $1.990$ or $1.993$, respectively. Removing pair regularization leaves the one-step prediction unchanged but raises the $\tau = 3$ error from $1.525$ to $1.923$. These comparisons support complementary representation and transition objectives without establishing exact disentanglement.

### 6.4 Clinical benchmarks

**Semi-synthetic MIMIC-III: treatment effects.** Clinical histories, binary vasopressor treatment, and SCM-generated outcomes support one-step effect evaluation (Liu et al., 2020; Melnychuk et al., 2022). SCM environments 0–3 provide 1,620 training, 200 validation, and 180 ID-test patients; environments 4–7 provide 2,000 OOD-test patients (Appendix D.1). This evaluates Stage 1 without diffusion rollout. We evaluate both individual and average effects because patient-level errors can cancel when averaged across a population.

Table 3: **Semi-synthetic MIMIC-III.** Next-step treatment-effect errors (mean ± SD, three seeds).

| Method | ID $\sqrt{\text{PEHE}}$ | OOD $\sqrt{\text{PEHE}}$ | OOD ATEErr |
|---|---|---|---|
| G-Net | 2.206±0.400 | 2.316±0.350 | 1.986±0.358 |
| CRN | 0.942±0.424 | 1.073±0.359 | 0.909±0.426 |
| CT | 0.742±0.140 | 0.735±0.049 | 0.482±0.126 |
| TARNet | 0.724±0.065 | 0.587±0.035 | 0.513±0.016 |
| G-Transf. | 0.411±0.043 | 0.513±0.073 | 0.235±0.024 |
| DR-CRN | 0.521±0.053 | 0.457±0.027 | 0.213±0.017 |
| CDM | 0.423±0.034 | 0.434±0.023 | 0.314±0.021 |
| **DECAF** | **0.372±0.092** | **0.314±0.004** | **0.157±0.017** |

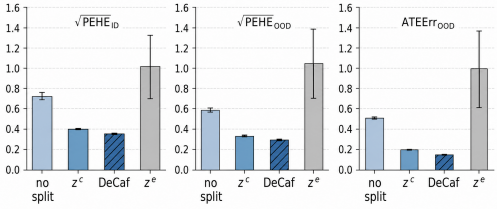


Figure 3: **Representation ablation.** ID $\sqrt{\text{PEHE}}$, OOD $\sqrt{\text{PEHE}}$, and OOD ATEErr for the unsplit, $z^c$, $z^e$, and full DECAF variants on semi-synthetic MIMIC-III. Lower is better.

DECAF has the lowest mean in each reported metric (Table 3). Compared with the best respective OOD baselines, $\sqrt{\text{PEHE}}$ decreases from $0.434$ (CDM) to $0.314$ ($27.6\%$) and ATEErr from $0.213$ (DR-CRN) to $0.157$ ($26.3\%$). Figure 3 also favors the full representation. We compare methods within each split because ID and OOD populations differ.

**Real MIMIC-IV: forecasting across care units.** We forecast $SpO_2$ given recorded future ventilation sequences (Johnson et al., 2023; Moukheiber et al., 2024a;b). Source units MICU, MICU/SICU, and SICU provide 2,144 training, 267 validation, and 271 ID-test stays. OOD comprises 686 CVICU and 732 TSICU stays, excluded from training and model selection (Appendix D.1).

Table 4: **Real MIMIC-IV factual forecasting on held-out care units.** We report $\text{nRMSE}_{\text{forecast}}(\tau) \downarrow$ as mean ± standard deviation over three seeds. Drift is $\text{nRMSE}_{\text{forecast}}(39)/\text{nRMSE}_{\text{forecast}}(1)$.

| Method | $\tau = 1$ | $\tau = 10$ | $\tau = 20$ | $\tau = 39$ | Drift |
|---|---|---|---|---|---|
| MSM-IPW | 1.556±0.005 | 5.659±0.007 | 8.039±0.022 | 9.881±0.026 | 6.35× |
| RMSN | 0.724±0.022 | 1.084±0.091 | 1.220±0.072 | 1.586±0.054 | 2.19× |
| CRN | 0.780±0.015 | 0.958±0.046 | 1.081±0.028 | 1.451±0.048 | 1.86× |
| G-Net | 0.802±0.026 | 1.136±0.051 | 1.300±0.066 | 1.649±0.033 | 2.06× |
| CT | 0.778±0.033 | 0.990±0.076 | 1.212±0.097 | 1.484±0.091 | 1.91× |
| DR-CRN | 0.749±0.012 | 0.861±0.003 | 0.871±0.021 | 1.190±0.005 | 1.59× |
| TE-CDE | 0.854±0.016 | 0.915±0.021 | 0.879±0.018 | 1.189±0.030 | **1.39**× |
| CDM | 0.702±0.023 | 0.892±0.034 | 0.946±0.023 | 1.542±0.081 | 2.20× |
| Persistence | 0.739±0.000 | 0.970±0.000 | 1.016±0.000 | 1.168±0.000 | 1.58× |
| **DECAF** | **0.634±0.000** | **0.785±0.004** | **0.824±0.008** | **0.985±0.008** | 1.55× |

DECAF has the lowest mean error at all reported horizons (Table 4); at $\tau = 39$, its error is $0.985$ versus $1.168$ for Persistence. Its drift ratio is $1.55\times$ versus $1.39\times$ for TE-CDE: TE-CDE has less relative growth but higher absolute errors. These results assess factual forecasting in held-out care units, not treatment-effect accuracy (Appendix D.5).

In the 24-hour patient-level analysis (Appendix D.1), nRMSE is $0.683$ for DECAF and $0.687$ for CRN. Each patient is weighted equally and only observed future $SpO_2$ hours are scored. The smaller gap highlights how window-averaged and per-horizon errors differ.

## 7 Conclusion and Discussion

DECAF separates outcome-relevant and environment-related representations and learns treatment-conditioned transitions in the outcome-relevant subspace. Under the appendix splits, it has the lowest mean error in 13 of 15 OOD Tumor Growth settings and the lowest reported OOD treatment-effect errors on semi-synthetic MIMIC-III. Held-out MIMIC-IV care units provide evidence for factual forecasting; mixed ID/OOD ablations assess the model components. Our theory states conditions for representation sufficiency, which the training objectives do not guarantee. Real EHR data lack oracle counterfactuals, so forecasting accuracy does not establish treatment-effect accuracy. Future work includes richer regimes and uncertainty-aware rollout for care planning and policy evaluation.

### AI USE STATEMENT

We used generative AI tools to assist with the implementation of parts of the proposed methods and to edit the manuscript for clarity and readability. All AI-assisted code was reviewed and tested by the authors to verify that it was consistent with the intended algorithms and experimental procedures. All AI-assisted edits to the manuscript were reviewed by the authors, and the authors take responsibility for the final content, implementation, experimental results, and claims presented in this work.

## REFERENCES

Farbod Alinezhad, Jianfei Cao, Gary J Young, and Brady Post. Causal diffusion models for counterfactual outcome distributions in longitudinal data. *arXiv preprint arXiv:2604.12992*, 2026. doi: 10.48550/arXiv.2604.12992. URL https://arxiv.org/abs/2604.12992.

Martin Arjovsky, Léon Bottou, Ishaan Gulrajani, and David Lopez-Paz. Invariant risk minimization. *arXiv preprint arXiv:1907.02893*, 2019.

Jeroen Berrevoets, Alicia Curth, Ioana Bica, Eoin McKinney, and Mihaela van der Schaar. Disentangled counterfactual recurrent networks for treatment effect inference over time. *arXiv preprint arXiv:2112.03811*, 2021.

Ioana Bica, Ahmed Alaa, and Mihaela Van Der Schaar. Time series deconfounder: Estimating treatment effects over time in the presence of hidden confounders. In *International conference on machine learning*, pp. 884–895. PMLR, 2020a.

Ioana Bica, Ahmed M Alaa, James Jordon, and Mihaela Van Der Schaar. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. *arXiv preprint arXiv:2002.04083*, 2020b.

Mouad El Bouchattaoui, Myriam Tami, Benoit Lepetit, and Paul-Henry Cournède. Causal dynamic variational autoencoder for counterfactual regression in longitudinal data. *arXiv preprint arXiv:2310.10559*, 2023.

Konstantinos Bousmalis, George Trigeorgis, Nathan Silberman, Dilip Krishnan, and Dumitru Erhan. Domain separation networks. *Advances in neural information processing systems*, 29, 2016.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PmLR, 2020.

Jiebin Chu, Yaoyun Zhang, Fei Huang, Luo Si, Songfang Huang, and Zhengxing Huang. Disentangled representation for sequential treatment effect estimation. *Computer Methods and Programs in Biomedicine*, 226:107175, 2022.

Anthony Cuturrufo, Fred Xu, Arvind Murari Vepa, David R Bakal, Veena K Ranganath, Yizhou Sun, and Wei Wang. Tidal: A temporal causal diffusion framework for visualizing knee osteoarthritis treatment outcomes. In *NeurIPS 2025 Workshop on Learning from Time Series for Health*.

Ann-Kathrin Dombrowski, Jan E Gerken, Klaus-Robert Müller, and Pan Kessel. Diffeomorphic counterfactuals with generative models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5):3257–3274, 2023.

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

Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A Wichmann. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11):665–673, 2020.

Changran Geng, Harald Paganetti, and Clemens Grassberger. Prediction of treatment response for combined chemo-and radiation therapy for non-small cell lung cancer patients using a biomathematical model. *Scientific reports*, 7(1):13542, 2017.

Miguel A Hernán and James M Robins. Causal inference, 2010.

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

Adrián Javaloy, Pablo Sánchez-Martín, and Isabel Valera. Causal normalizing flows: from theory to practice. *Advances in Neural Information Processing Systems*, 36:58833–58864, 2023.

Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *International conference on machine learning*, pp. 3020–3029. PMLR, 2016.

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

Nathan Kallus, Xiaojie Mao, and Madeleine Udell. Causal inference with noisy and missing covariates via matrix factorization. *Advances in neural information processing systems*, 31, 2018.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Rui Li, Zach Shahn, Jun Li, Mingyu Lu, Prithwish Chakraborty, Daby Sow, Mohamed Ghalwash, and Li-wei H Lehman. G-net: a deep learning approach to g-computation for counterfactual outcome prediction under dynamic treatment regimes. *arXiv preprint arXiv:2003.10551*, 2020.

Bryan Lim. Forecasting treatment responses over time using recurrent marginal structural networks. *Advances in neural information processing systems*, 31, 2018.

Ruoqi Liu, Changchang Yin, and Ping Zhang. Estimating individual treatment effects with time-varying confounders. In *2020 IEEE international conference on data mining (ICDM)*, pp. 382–391. IEEE, 2020.

Ruoqi Liu, Katherine M Hunold, Jeffrey M Caterino, and Ping Zhang. Estimating treatment effects for time-to-treatment antibiotic stewardship in sepsis. *Nature machine intelligence*, 5(4):421–431, 2023.

Francesco Locatello, Ben Poole, Gunnar Rätsch, Bernhard Schölkopf, Olivier Bachem, and Michael Tschannen. Weakly-supervised disentanglement without compromises. In *International conference on machine learning*, pp. 6348–6359. PMLR, 2020.

Christos Louizos, Uri Shalit, Joris M Mooij, David Sontag, Richard Zemel, and Max Welling. Causal effect inference with deep latent-variable models. *Advances in neural information processing systems*, 30, 2017.

Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. Causal transformer for estimating counterfactual outcomes. In *International conference on machine learning*, pp. 15293–15329. PMLR, 2022.

M Moukheiber, L Moukheiber, D Moukheiber, S Hao, LA Celi, and HC Lee. A temporal dataset for respiratory support in critically ill patients (version 1.0. 0). *PhysioNet*, 2024a.

Mira Moukheiber, Lama Moukheiber, Dana Moukheiber, and Hyung-Chul Lee. Unmasking societal biases in respiratory support for icu patients through social determinants of health. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pp. 7421–7429, 2024b. doi: 10.24963/ijcai.2024/821. URL https://www.ijcai.org/proceedings/2024/821.

Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

Jonas Peters, Peter Bühlmann, and Nicolai Meinshausen. Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(5):947–1012, 2016.

Vishal Purohit. Counterfactual outcome prediction using structured state space model. *arXiv preprint arXiv:2305.09207*, 2023.

Zhaozhi Qian, Yao Zhang, Ioana Bica, Angela Wood, and Mihaela van der Schaar. Synctwin: Treatment effect estimation with longitudinal outcomes. *Advances in Neural Information Processing Systems*, 34:3178–3190, 2021.

Kashif Rasul, Calvin Seward, Ingmar Schuster, and Roland Vollgraf. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. In *International conference on machine learning*, pp. 8857–8868. PMLR, 2021.

Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pp. 1530–1538. PMLR, 2015.

James Robins. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical modelling*, 7 (9-12):1393–1512, 1986.

James M Robins, Miguel Angel Hernan, and Babette Brumback. Marginal structural models and causal inference in epidemiology, 2000.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

Shiori Sagawa, Pang Wei Koh, Tatsunori B Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. *arXiv preprint arXiv:1911.08731*, 2019.

Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021.

Nabeel Seedat, Fergus Imrie, Alexis Bellot, Zhaozhi Qian, and Mihaela van der Schaar. Continuous-time modeling of counterfactual outcomes using neural controlled differential equations. *arXiv preprint arXiv:2206.08311*, 2022.

Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pp. 3076–3085. PMLR, 2017.

Kihyuk Sohn, Honglak Lee, and Xinchen Yan. Learning structured output representation using deep conditional generative models. In *Advances in Neural Information Processing Systems*, volume 28, 2015.

Hossein Soleimani, Adarsh Subbaswamy, and Suchi Saria. Treatment-response models for counterfactual reasoning with continuous-time, continuous-valued interventions. 2017.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. URL `https://arxiv.org/abs/2010.02502`.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Xin Wang, Shengfei Lyu, Kangyang Luo, Lishan Yang, Huanhuan Chen, and Chunyan Miao. Enhancing counterfactual estimation: a focus on temporal treatments. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 6442–6450, 2025.

Yixin Wang and David M Blei. The blessings of multiple causes. *Journal of the American Statistical Association*, 114(528):1574–1596, 2019.

Zhendong Wang, Ioanna Miliou, Isak Samsten, and Panagiotis Papapetrou. Counterfactual explanations for time series forecasting. In *2023 IEEE International Conference on Data Mining (ICDM)*, pp. 1391–1396. IEEE, 2023.

Dongze Wu, David I Inouye, and Yao Xie. Po-flow: Flow-based generative models for sampling potential outcomes and counterfactuals. *arXiv preprint arXiv:2505.16051*, 2025a.

Dongze Wu, Feng Qiu, and Yao Xie. Doflow: Causal generative flows for interventional and counterfactual time-series prediction. *arXiv preprint arXiv:2511.02137*, 2025b.

Shenghao Wu, Wenbin Zhou, Minshuo Chen, and Shixiang Zhu. Counterfactual generative models for time-varying treatments. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3402–3413, 2024.

Hong Xiong, Feng Wu, Leon Deng, Megan Su, Zach Shahn, and Li-wei H Lehman. G-transformer: Counterfactual outcome prediction under dynamic and time-varying treatment regimes. volume 252, pp. https–proceedings, 2024.

Liuyi Yao, Sheng Li, Yaliang Li, Mengdi Huai, Jing Gao, and Aidong Zhang. Representation learning for treatment effect estimation from observational data. *Advances in neural information processing systems*, 31, 2018.

Jinsung Yoon, James Jordon, and Mihaela Van Der Schaar. Ganite: Estimation of individualized treatment effects using generative adversarial nets. In *International conference on learning representations*, 2018.

Jiaxuan Zhang, Emadeldeen Eldele, Yang Wang, Xiaoli Li, Jiye Liang, et al. Counterfactual contrastive learning with normalizing flows for robust treatment effect estimation. In *Forty-second International Conference on Machine Learning*, 2025.

# A Notation and Training Algorithms

DeCaf combines a history encoder, treatment-specific outcome heads, and a latent transition model. Stage 1 learns the encoder and outcome heads; Stage 2 fits the transition model with the Stage 1 parameters fixed. This section defines the prediction targets and training objectives, followed by the forecasting procedure. Table 5 collects the notation. Bold symbols denote vectors or matrices, calligraphic symbols denote sets, and hats denote estimates. The superscripts $c$ and $e$ identify the outcome-relevant and environment-related branches. The plain symbols $z_t^c, z_t^e$ are used interchangeably with $\mathbf{z}_t^c, \mathbf{z}_t^e$.

**Treatments and observed histories.** Treatments take values in a finite set $\mathcal{A}$, with $|\mathcal{A}| \geq 2$. We write $\mathrm{a}_t$ for the observed assignment and $a$ for a specified treatment value. Binary treatments use $\mathcal{A} = \{0, 1\}$; multiple binary treatments are represented by their joint categories. In Tumor Growth, radiotherapy and chemotherapy give $\mathcal{A} = \{(0,0), (1,0), (0,1), (1,1)\}$. Each category has an outcome head and serves as an input to the transition model.

The history available before treatment at time $t$ is

$$\mathbb{H}_t = (\mathbf{X}_t, \boldsymbol{a}_{t-1}, \boldsymbol{y}_t), \qquad \mathbf{X}_t = [\mathbf{x}_1^\top; \ldots; \mathbf{x}_t^\top], \tag{4}$$

where $\boldsymbol{a}_{t-1} = (\mathrm{a}_1, \ldots, \mathrm{a}_{t-1})$ and $\boldsymbol{y}_t = (\mathrm{y}_1, \ldots, \mathrm{y}_t)$. Thus $\mathrm{y}_t$ is observed before $\mathrm{a}_t$ is assigned, and $\mathrm{y}_{t+1}$ is the next-step outcome. The recurrent encoder receives $(\mathbf{x}_s, \mathrm{a}_{s-1}, \mathrm{y}_s)$ at each $s \leq t$, with a fixed initialization token for $\mathrm{a}_0$. All inputs are available before the corresponding treatment decision. Training uses decision times with an observed next-step outcome; Stage 2 also requires the adjacent history $\mathbb{H}_{t+1}$.

**Prediction targets.** Let $\mathrm{y}_{t+1}^a$ denote the potential outcome under treatment $a$ at time $t$, with past treatments unchanged. For a future plan $\boldsymbol{a}'_{t:t+\tau-1} = (a'_t, \ldots, a'_{t+\tau-1})$, the outcome $\mathrm{y}_{t+r+1}^{\boldsymbol{a}'}$, $0 \leq r < \tau$, corresponds to its prefix $(a'_t, \ldots, a'_{t+r})$. We estimate the conditional mean

$$\mu_{t+r+1}^{\boldsymbol{a}'}(\mathbb{H}_t) = \mathbb{E}\left[\mathrm{y}_{t+r+1}^{\boldsymbol{a}'} \mid \mathbb{H}_t\right]. \tag{5}$$

For binary treatment, the one-step effect is $\tau_t(\mathbb{H}_t) = \mathbb{E}[\mathrm{y}_{t+1}^1 - \mathrm{y}_{t+1}^0 \mid \mathbb{H}_t]$. Throughout the paper, individualized treatment effects refer to these history-conditional mean contrasts. We use $\tau$ for the forecast horizon, $r$ for a rollout offset, $k$ for a diffusion step, and $\kappa_{\mathrm{con}}$ for the contrastive temperature.

Table 5: Notation used in the model, algorithms, and evaluation.

| Symbol | Meaning |
|---|---|
| *Indices, sets, and dimensions* | |
| $N$ | Number of observational units (longitudinal trajectories). |
| $T^{(n)}, T$ | Observed length of unit $n$ and the maximum observed trajectory length. |
| $t, s$ | Clinical decision-time indices. |
| $\tau, r$ | Forecast horizon and rollout offset, with $0 \le r < \tau$. |
| $\mathcal{A}, \mathcal{E}, \mathcal{T}$ | Treatment alphabet, environment-label set, and evaluation-horizon set. |
| $d$ | Covariate dimension; observable embedding dimension in the linear analysis. |
| $d_c, d_e$ | Dimensions of the outcome-relevant and environment-related representation branches. |
| $K, k$ | Number of diffusion noising steps and diffusion-step index, $k \in \{1, \dots, K\}$. |
| $\mathcal{B}, B$ | A minibatch and its size; $i, j$ index its examples. |
| $M, m$ | Number of Monte Carlo rollouts and rollout index, $m \in \{1, \dots, M\}$. |
| *Observed data and intervention targets* | |
| $\mathbf{x}_t \in \mathbb{R}^d$ | Covariates available at decision time $t$. |
| $\mathbf{X}_t \in \mathbb{R}^{t \times d}$ | Matrix with rows $\mathbf{x}_1^\top, \dots, \mathbf{x}_t^\top$. |
| $\mathrm{a}_t \in \mathcal{A}$ | Observed treatment assignment at time $t$. |
| $\boldsymbol{a}_t$ | Observed treatment sequence $(\mathrm{a}_1, \dots, \mathrm{a}_t) \in \mathcal{A}^t$. |
| $\mathrm{y}_t, \boldsymbol{y}_t$ | Scalar outcome available at time $t$ and outcome history $(\mathrm{y}_1, \dots, \mathrm{y}_t)$. |
| $\mathbb{H}_t$ | Pre-treatment history in Eq. equation 4; it excludes $\mathrm{a}_t$ and future measurements. |
| $\mathrm{e} \in \mathcal{E}$ | Trajectory-level environment label used for the training objectives. |
| $\mathcal{D}$ | Dataset of trajectories $(\mathbf{X}^{(n)}, \boldsymbol{a}^{(n)}, \boldsymbol{y}^{(n)}, \mathrm{e}^{(n)})$, $n = 1, \dots, N$. |
| $\mathrm{y}_{t+1}^a$ | Potential next-step outcome under treatment $a$ at time $t$, with the observed past unchanged. |
| $\boldsymbol{a}'_{t:t+\tau-1}$ | Fixed future treatment plan $(a'_t, \dots, a'_{t+\tau-1}) \in \mathcal{A}^\tau$. |
| $\mu_{t+r+1}^{\boldsymbol{a}'}(\mathbb{H}_t)$ | Conditional mean potential outcome under the relevant plan prefix, defined in Eq. equation 5. |
| $\widehat{\mu}_{t+r+1}^{\boldsymbol{a}'}(\mathbb{H}_t)$ | Model estimate of this conditional mean. The point-prediction notation $\widehat{\mathrm{y}}_{t+r+1}^{\boldsymbol{a}'}$ refers to the same estimate. |
| $\tau_t(\mathbb{H}_t), \widehat{\tau}_t$ | Binary-treatment conditional mean effect and its estimate, respectively. |
| *Representations and prediction models* | |
| $\mathbf{z}_t = (\mathbf{z}_t^c, \mathbf{z}_t^e)$ | Decomposed representation of $\mathbb{H}_t$. |
| $\mathbf{z}_t^c \in \mathbb{R}^{d_c}$ | Outcome-relevant branch used by the outcome heads and latent transition model. |
| $\mathbf{z}_t^e \in \mathbb{R}^{d_e}$ | Auxiliary branch trained to retain environment-predictive information. |
| $f_{\theta_f}, f_{\theta_f}^c$ | History encoder and its outcome-relevant output, respectively. |
| $g_a, \theta_g$ | Outcome head for treatment $a$ and the collection of parameters of all outcome heads. |
| $g(\mathbf{z}_t^c)$ | Vector $(g_a(\mathbf{z}_t^c))_{a \in \mathcal{A}}$ in a fixed treatment-category order. |
| $h_{\mathrm{env}}^c, h_{\mathrm{env}}^e$ | Environment classifiers for the outcome-relevant and environment-related branches. |
| $h_{\mathrm{suf},a}^c$ | Treatment-specific outcome probe using $\mathbf{z}_t^c$ alone. |
| $h_{\mathrm{suf},a}^{\mathrm{aug}}$ | Treatment-specific augmented probe using $(\mathbf{z}_t^c, \mathbb{H}_t)$. |
| $h, q$ | Projection heads in the same-history pair objective. |
| *Diffusion and sampling* | |
| $\beta_k, \alpha_k, \bar{\alpha}_k$ | Noising schedule and coefficients: $0 < \beta_k < 1$, $\alpha_k = 1 - \beta_k$, $\bar{\alpha}_k = \prod_{j=1}^k \alpha_j$. |
| $\mathbf{z}_{t+1,k}^c$ | Noised factual next-step latent at diffusion step $k$; $\mathbf{z}_{t+1,0}^c = \mathbf{z}_{t+1}^c$. |
| $\epsilon$ | Independent standard Gaussian training noise in $\mathbb{R}^{d_c}$. |
| $\epsilon_\psi$ | Conditional noise-prediction network with parameters $\psi$. |

*Continued on next page*

Table 5: continued

| Symbol | Meaning |
|---|---|
| $G_\psi(\mathbf{z}, a; \xi)$ | Differentiable reverse sampler, including its initialization and sampling updates; $\xi$ collects its parameter-independent random draws. |
| $p_\psi(\cdot \mid \mathbf{z}, a)$ | Output distribution induced by the specified sampler $G_\psi(\mathbf{z}, a; \xi)$. |
| $\widetilde{\mathbf{z}}^c_{t+1}$ | Generated next-step latent under an alternative treatment; also written $\mathbf{z}^c_{t+1,\mathrm{cf}}$. |
| $\mathbf{z}_s^{c,(m)}$ | Outcome-relevant state of Monte Carlo rollout $m$ at clinical time $s$. |
| *Optimization quantities* | |
| $\widehat{\mathbb{E}}_{\mathcal{B}}$ | Arithmetic mean over the valid examples in minibatch $\mathcal{B}$. |
| $\mathrm{sg}(\cdot)$ | Stop-gradient operator: unchanged value in the forward pass and zero derivative in the backward pass. |
| $\mathrm{GRL}(\cdot)$ | Gradient reversal: unchanged forward value and reversed input gradient. |
| $\ell_{\mathrm{CE}}$ | Cross-entropy classification loss. |
| $\mathcal{L}_{\mathrm{reg}}$ | Mean squared prediction error over the relevant examples. |
| $\mathcal{L}_y$ | Factual next-step outcome loss. |
| $\mathcal{L}^c_{\mathrm{env}}, \mathcal{L}^e_{\mathrm{env}}$ | Positive cross-entropy losses of the two environment classifiers. |
| $\mathcal{L}_{\mathrm{env}}$ | When used with GRL, the sum of the two positive classification losses with GRL on the $c$ branch; this is a computational objective with parameter-specific gradients. |
| $\ell_c, \ell_{\mathrm{aug}}$ | Squared losses of the outcome-relevant and augmented sufficiency probes. |
| $\mathcal{L}_{\mathrm{suf}}$ | Positive-part probe-gap penalty in Eq. equation 10. |
| $J_f$ | Encoder-side objective with explicit adversarial signs in Eq. equation 11. |
| $\mathcal{L}_{\mathrm{diff}}$ | Factual-transition diffusion noise-prediction loss. |
| $\mathcal{L}_{\mathrm{con}}$ | Same-history pair contrastive loss. |
| $\lambda_{\mathrm{env}}, \lambda_{\mathrm{suf}}, \lambda_{\mathrm{con}}$ | Nonnegative weights for the corresponding objective terms. |
| $\delta, \kappa_{\mathrm{con}}$ | Nonnegative probe-gap slack and strictly positive contrastive temperature. |
| $\mathrm{sim}(\cdot, \cdot)$ | Cosine similarity, with a fixed numerical stabilizer in its denominator. |
| *Stylized linear-Gaussian analysis* | |
| $\mathbf{s}_t, \mathbf{u}_t$ | Predictive and nuisance states in the linear-Gaussian analysis. |
| $\mu_e, \Sigma_U$ | Environment-specific nuisance mean and common nuisance covariance in that model. |
| $\mathbf{h}_t = \mathbf{A}\mathbf{s}_t + \mathbf{B}\mathbf{u}_t$ | Observable history embedding in the linear analysis. |
| $\mathbf{A}, \mathbf{B}$ | Mixing matrices of dimensions $d \times d_c$ and $d \times d_e$; the recovery result assumes their concatenation has full column rank. |
| $\mathbf{W} \in \mathbb{R}^{d \times d_c}$ | Linear representation matrix, with $\mathbf{z}^c_t = \mathbf{W}^\top \mathbf{h}_t$. |
| $\beta_a \in \mathbb{R}^{d_c}$ | Treatment-specific outcome coefficient; distinct from the scalar diffusion coefficient $\beta_k$. |
| $R \in \mathbb{R}^{d_c \times d_c}$ | Invertible change of basis in the full-state recovery result. |
| *Evaluation metrics* | |
| $\sqrt{\mathrm{PEHE}}$ | Root mean squared error of the estimated treatment-effect contrast against its oracle value. |
| ATEErr | Absolute error in the average treatment-effect contrast over the evaluation population. |
| $\mathrm{nRMSE}_{\mathrm{cf}}(\tau)$ | Normalized prediction error at horizon $\tau$ for specified intervention plans when oracle outcomes are available. |
| $\mathrm{nRMSE}_{\mathrm{forecast}}(\tau)$ | Normalized factual forecasting error at horizon $\tau$. |

### A.1 Stage 1: Decomposed Representation Learning

The encoder maps each eligible history to $(\mathbf{z}^c_t, \mathbf{z}^e_t) = f_{\theta_f}(\mathbb{H}_t)$. The outcome heads use $\mathbf{z}^c_t$ to predict treatment-specific outcomes. Two auxiliary objectives guide the allocation of information between the branches: environment classification and outcome sufficiency. Below, the unit index is suppressed and $\widehat{\mathbb{E}}_{\mathcal{B}}$ denotes the mean over a minibatch of valid unit–time examples.

**Outcome prediction and environment separation.** The factual prediction loss and the two environment-classification losses are

$$\mathcal{L}_y = \widehat{\mathbb{E}}_{\mathcal{B}}\left[(\mathrm{y}_{t+1} - g_{\mathrm{a}_t}(\mathbf{z}_t^c))^2\right], \tag{6}$$

$$\mathcal{L}_{\mathrm{env}}^c = \widehat{\mathbb{E}}_{\mathcal{B}}\ell_{\mathrm{CE}}(\mathrm{e}, h_{\mathrm{env}}^c(\mathbf{z}_t^c)), \qquad \mathcal{L}_{\mathrm{env}}^e = \widehat{\mathbb{E}}_{\mathcal{B}}\ell_{\mathrm{CE}}(\mathrm{e}, h_{\mathrm{env}}^e(\mathbf{z}_t^e)). \tag{7}$$

Both classifiers minimize cross-entropy using detached encoder outputs. Their objectives have different signs in the encoder update: the encoder maximizes $\mathcal{L}_{\mathrm{env}}^c$ and minimizes $\mathcal{L}_{\mathrm{env}}^e$. This adversarial update discourages environment prediction from $\mathbf{z}_t^c$, while the second classifier encourages $\mathbf{z}_t^e$ to retain environment-related information.

**Outcome-sufficiency regularization.** We compare two treatment-specific probes. The first predicts the outcome from $\mathbf{z}_t^c$; the second additionally receives the full history. Their squared losses are

$$\ell_c = \widehat{\mathbb{E}}_{\mathcal{B}}\left[(\mathrm{y}_{t+1} - h_{\mathrm{suf},\mathrm{a}_t}^c(\mathbf{z}_t^c))^2\right], \tag{8}$$

$$\ell_{\mathrm{aug}} = \widehat{\mathbb{E}}_{\mathcal{B}}\left[(\mathrm{y}_{t+1} - h_{\mathrm{suf},\mathrm{a}_t}^{\mathrm{aug}}(\mathbf{z}_t^c, \mathbb{H}_t))^2\right]. \tag{9}$$

At each iteration, the probes minimize $\ell_c + \ell_{\mathrm{aug}}$ with the encoder output detached. We then fix their parameters and recompute both losses for the encoder update. The regularizer penalizes the excess loss of the representation-only probe beyond a slack $\delta$:

$$\mathcal{L}_{\mathrm{suf}} = [\ell_c - \mathrm{sg}(\ell_{\mathrm{aug}}) - \delta]_+, \qquad [u]_+ = \max(0, u), \qquad \delta \geq 0. \tag{10}$$

The stop-gradient operator holds the augmented loss fixed as a reference during this update. Gradients reach the encoder through the input of the representation-only probe. Both losses are minibatch means, and the reference is refreshed at the next iteration.

**Alternating optimization.** With the auxiliary models fixed, the encoder minimizes

$$J_f = \mathcal{L}_y + \lambda_{\mathrm{env}}(\mathcal{L}_{\mathrm{env}}^e - \mathcal{L}_{\mathrm{env}}^c) + \lambda_{\mathrm{suf}}\mathcal{L}_{\mathrm{suf}}. \tag{11}$$

The outcome heads minimize $\mathcal{L}_y$. Equivalently, the adversarial encoder gradient can be implemented with a gradient reversal layer before $h_{\mathrm{env}}^c$ and a positive classification loss, in place of the explicit negative sign in Eq. equation 11.

Algorithm 1 gives the update sequence. The auxiliary parameter groups are $\theta_{\mathrm{env}}$ and $\theta_{\mathrm{suf}}$, and $\mathsf{AdamStep}(\vartheta, d)$ applies one Adam update with gradient $d$ and the parameter group's optimizer state. The operator $\mathsf{Batch}$ samples valid unit–time examples; a superscript $+$ denotes the next observed time point. We write $Z^c = (z_i^c)_{i=1}^B$, $Z^e = (z_i^e)_{i=1}^B$, and $\mathbb{H}_{\mathcal{B}} = (\mathbb{H}_i)_{i=1}^B$. $\mathsf{Losses}$ evaluates Eqs. equation 6–equation 9 on the current minibatch. Stopped parameter copies fix the auxiliary models while preserving derivatives with respect to their inputs. The two stages use $N_1$ and $N_2$ training iterations, respectively.

The probe gap compares predictive performance with and without access to the full history under the observed treatment distribution. Its dependence on probe capacity and optimization makes it an empirical criterion for sufficiency. Population conditions for potential-outcome identification are given in Appendix H; the training objectives alone do not establish a unique decomposition or independence of the two branches.

### A.2 Stage 2: Treatment-Conditional Latent Transition Learning

Stage 2 learns treatment-conditional transitions from adjacent factual histories, with the encoder and outcome heads fixed. Each history pair provides $(\mathbf{z}_t^c, \mathrm{a}_t, \mathbf{z}_{t+1}^c)$ for fitting the transition law $p_\psi(\cdot \mid \mathbf{z}_t^c, \mathrm{a}_t)$. The training objective combines diffusion on factual transitions with a contrastive term that pairs factual and generated transitions from the same history.

**Factual transition modeling.** For each example, independently sample $k \sim \mathrm{Unif}\{1, \ldots, K\}$ and $\epsilon \sim \mathcal{N}(0, \mathbf{I}_{d_c})$, and construct

$$\mathbf{z}_{t+1,k}^c = \sqrt{\bar{\alpha}_k}\mathbf{z}_{t+1}^c + \sqrt{1 - \bar{\alpha}_k}\epsilon, \qquad \bar{\alpha}_k = \prod_{j=1}^{k}(1 - \beta_j). \tag{12}$$

**Algorithm 1** Stage 1: Alternating optimization of decomposed representations

**Input:** Training data $\mathcal{D}$; initialized parameters $\theta_f, \theta_g, \theta_{\text{env}}, \theta_{\text{suf}}$; batch size $B$; iterations $N_1$; weights $\lambda_{\text{env}}, \lambda_{\text{suf}}$; slack $\delta$.
**Output:** Trained encoder $f_{\theta_f}$ and outcome heads $\{g_a\}_{a\in\mathcal{A}}$.

**for** $j = 1, \dots, N_1$ **do**
  $\mathcal{B} = \{(\mathbb{H}_i, a_i, y_i^+, e_i)\}_{i=1}^B \leftarrow \mathsf{Batch}(\mathcal{D}, B)$
  $(Z^c, Z^e) \leftarrow f_{\theta_f}(\mathbb{H}_{\mathcal{B}})$
  $(\bar{Z}^c, \bar{Z}^e) \leftarrow \text{sg}(Z^c, Z^e)$
  $d_{\text{env}} \leftarrow \nabla_{\theta_{\text{env}}}[\mathcal{L}^c_{\text{env}}(\bar{Z}^c) + \mathcal{L}^e_{\text{env}}(\bar{Z}^e)]$
  $\theta_{\text{env}} \leftarrow \mathsf{AdamStep}(\theta_{\text{env}}, d_{\text{env}})$
  $d_{\text{suf}} \leftarrow \nabla_{\theta_{\text{suf}}}[\ell_c(\bar{Z}^c) + \ell_{\text{aug}}(\bar{Z}^c, \mathbb{H}_{\mathcal{B}})]$
  $\theta_{\text{suf}} \leftarrow \mathsf{AdamStep}(\theta_{\text{suf}}, d_{\text{suf}})$
  $(\bar{\theta}_{\text{env}}, \bar{\theta}_{\text{suf}}) \leftarrow \text{sg}(\theta_{\text{env}}, \theta_{\text{suf}})$
  $(\mathcal{L}_y, \mathcal{L}^c_{\text{env}}, \mathcal{L}^e_{\text{env}}, \ell_c, \ell_{\text{aug}}) \leftarrow$
   $\mathsf{Losses}(\mathcal{B}, Z^c, Z^e; \theta_g, \bar{\theta}_{\text{env}}, \bar{\theta}_{\text{suf}})$
  $\mathcal{L}_{\text{suf}} \leftarrow [\ell_c - \text{sg}(\ell_{\text{aug}}) - \delta]_+$
  $J_f \leftarrow \mathcal{L}_y + \lambda_{\text{env}}(\mathcal{L}^e_{\text{env}} - \mathcal{L}^c_{\text{env}}) + \lambda_{\text{suf}}\mathcal{L}_{\text{suf}}$
  $(d_f, d_g) \leftarrow (\nabla_{\theta_f} J_f, \nabla_{\theta_g} \mathcal{L}_y)$
  $\theta_f \leftarrow \mathsf{AdamStep}(\theta_f, d_f)$
  $\theta_g \leftarrow \mathsf{AdamStep}(\theta_g, d_g)$
**end for**
**return** $f_{\theta_f}, \{g_a\}_{a\in\mathcal{A}}$

The noise-prediction objective is

$$\mathcal{L}_{\text{diff}} = \widehat{\mathbb{E}}_{\mathcal{B}}\, \mathbb{E}_{k,\epsilon}\left[\left\|\epsilon - \epsilon_\psi(\mathbf{z}^c_{t+1,k}, k \mid \mathbf{z}^c_t, \mathrm{a}_t)\right\|_2^2\right]. \tag{13}$$

Each stochastic update uses one diffusion step and one noise draw per example. The next latent $\mathbf{z}^c_{t+1}$ is the denoising target, while the conditioning variables are the current latent and factual treatment.

**Transitions under alternative treatments.** Let $G_\psi(\mathbf{z}, a; \xi)$ denote the reverse sampler, including its initialization and timestep sequence. Its random inputs $\xi$ have a distribution independent of $\psi$, and its output distribution defines $p_\psi(\cdot \mid \mathbf{z}, a)$. We sample an alternative treatment $a'_t$ uniformly from $\mathcal{A} \setminus \{\mathrm{a}_t\}$ and generate

$$\widetilde{\mathbf{z}}^c_{t+1} = G_\psi(\mathbf{z}^c_t, a'_t; \xi). \tag{14}$$

Encoder outputs remain detached, while gradients through $G_\psi$ are retained so that the pair objective can update the transition model.

**Contrastive pair objective.** A minibatch contains one transition from each of $B$ distinct units. For example $i$, the factual and generated next states are projected as

$$u_i = h([\mathbf{z}^{c,(i)}_{t_i+1}; \mathbf{z}^{e,(i)}_{t_i}]), \qquad v_i = q([\widetilde{\mathbf{z}}^{c,(i)}_{t_i+1}; \mathbf{z}^{e,(i)}_{t_i}]), \tag{15}$$

where $(i)$ indexes the example and $[\cdot;\cdot]$ denotes concatenation. The pair $(u_i, v_i)$ shares the pre-treatment history and environment-related representation, but corresponds to different current treatments. Pairs with different indices come from different units. The contrastive loss is

$$\mathcal{L}_{\text{con}} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\text{sim}(u_i, v_i)/\kappa_{\text{con}})}{\sum_{j=1}^{B} \exp(\text{sim}(u_i, v_j)/\kappa_{\text{con}})}, \qquad \kappa_{\text{con}} > 0. \tag{16}$$

The denoiser and projection heads jointly minimize

$$\mathcal{L}_{\text{stage2}} = \mathcal{L}_{\text{diff}} + \lambda_{\text{con}}\mathcal{L}_{\text{con}}. \tag{17}$$

The factual loss fits the observed transition distribution. The pair loss encourages the projections to retain structure shared across treatments, without equating the next-step latents or their outcome predictions. Since both projections also receive $\mathbf{z}^e_t$, alignment may use this shared input; the pair loss alone does not establish the accuracy of the generated transition or its treatment-effect contrast.

**Algorithm 2** Stage 2: Treatment-conditional diffusion with pair regularization

**Input:** Training data $\mathcal{D}$; frozen encoder $f_{\theta_f}$; initialized denoiser $\epsilon_\psi$ and projections $h_{\theta_h}, q_{\theta_q}$; sampler $G_\psi$; noise schedule $\{\bar{\alpha}_k\}_{k=1}^K$; $B, N_2, \lambda_{\text{con}}, \kappa_{\text{con}}$.
**Output:** Fitted transition sampler $G_\psi$.

$\vartheta_2 \leftarrow (\psi, \theta_h, \theta_q)$
**for** $j = 1, \ldots, N_2$ **do**
  $\mathcal{B} = \{(\mathbb{H}_i, a_i, \mathbb{H}_i^+)\}_{i=1}^B \leftarrow \mathsf{Transitions}(\mathcal{D}, B)$
  **for** $i = 1, \ldots, B$ **do**
    $(z_i^c, z_i^e) \leftarrow \text{sg}(f_{\theta_f}(\mathbb{H}_i))$; $z_i^{c,+} \leftarrow \text{sg}(f_{\theta_f}^c(\mathbb{H}_i^+))$
    $k_i \sim \text{Unif}\{1, \ldots, K\}$; $\epsilon_i \sim \mathcal{N}(0, \mathbf{I}_{d_c})$
    $w_i \leftarrow \sqrt{\bar{\alpha}_{k_i}} z_i^{c,+} + \sqrt{1 - \bar{\alpha}_{k_i}} \epsilon_i$
    $a_i' \sim \text{Unif}(\mathcal{A} \setminus \{a_i\})$; $\xi_i \sim P_\xi$
    $\widetilde{z}_i^{c,+} \leftarrow G_\psi(z_i^c, a_i'; \xi_i)$
    $u_i \leftarrow h_{\theta_h}([z_i^{c,+}; z_i^e])$; $v_i \leftarrow q_{\theta_q}([\widetilde{z}_i^{c,+}; z_i^e])$
  **end for**
  $\mathcal{L}_{\text{diff}} \leftarrow B^{-1} \sum_{i=1}^B \|\epsilon_i - \epsilon_\psi(w_i, k_i \mid z_i^c, a_i)\|_2^2$
  $\mathcal{L}_{\text{con}} \leftarrow -B^{-1} \sum_{i=1}^B \log \dfrac{\exp(\text{sim}(u_i, v_i)/\kappa_{\text{con}})}{\sum_{l=1}^B \exp(\text{sim}(u_i, v_l)/\kappa_{\text{con}})}$
  $d_2 \leftarrow \nabla_{\vartheta_2}(\mathcal{L}_{\text{diff}} + \lambda_{\text{con}} \mathcal{L}_{\text{con}})$
  $\vartheta_2 \leftarrow \mathsf{AdamStep}(\vartheta_2, d_2)$
**end for**
**return** $G_\psi$

Algorithm 2 summarizes training. Transitions draws one adjacent transition per patient, $\vartheta_2 = (\psi, \theta_h, \theta_q)$ collects the trainable parameters, and $P_\xi$ is the sampler's parameter-independent noise law. Random draws are independent unless a shared quantity is specified. The factual branch is conditioned on $a_i$ and the generated branch on $a_i'$.

### A.3 Counterfactual Rollout

For a history $\mathbb{H}_t$ and a prescribed treatment plan $\boldsymbol{a}'_{t:t+\tau-1}$, initialize $M$ rollout paths at the encoded state:

$$\mathbf{z}_t^{c,(m)} = f_{\theta_f}^c(\mathbb{H}_t), \qquad m = 1, \ldots, M. \tag{18}$$

At decision time $s$, the outcome head predicts $\mathrm{y}_{s+1}$ from the current latent. A transition is sampled when another forecast step is required:

$$\widehat{\mu}_{s+1}^{\boldsymbol{a}'}(\mathbb{H}_t) = \frac{1}{M} \sum_{m=1}^M g_{a_s'}(\mathbf{z}_s^{c,(m)}), \tag{19}$$

$$\mathbf{z}_{s+1}^{c,(m)} = G_\psi(\mathbf{z}_s^{c,(m)}, a_s'; \xi_{s,m}), \qquad s < t + \tau - 1. \tag{20}$$

Noise draws are independent across paths and clinical transitions. Each path propagates sampled latent states directly from its initial history, without access to future observations. Only $\mathbf{z}^c$ and the prescribed treatments enter the outcome heads and transition sampler; the environment label and $\mathbf{z}^e$ are used during training only.

Algorithm 3 abbreviates $\mathbf{z}_s^{c,(m)}$ as $z_s^{(m)}$. All forecasts remain conditional on the initial history $\mathbb{H}_t$ and are evaluated under the corresponding prefix of the supplied plan.

**Treatment-effect estimates and interpretation.** At $\tau = 1$, the forecast is $g_{a_t'}(f_{\theta_f}^c(\mathbb{H}_t))$. For binary treatment, the estimated effect is

$$\widehat{\tau}_t(\mathbb{H}_t) = g_1(f_{\theta_f}^c(\mathbb{H}_t)) - g_0(f_{\theta_f}^c(\mathbb{H}_t)). \tag{21}$$

Longer-horizon effects are obtained by contrasting the mean forecasts under two treatment plans. A horizon-$\tau$ forecast uses $\tau - 1$ latent transitions. Averaging the head outputs approximates a conditional mean under the fitted model; an additional outcome-noise model would be needed to

**Algorithm 3** Monte Carlo forecasting under a fixed treatment plan

**Input:** Pre-treatment history $\mathbb{H}_t$; plan $\boldsymbol{a}'_{t:t+\tau-1}$; encoder $f_{\theta_f}$; outcome heads $\{g_a\}_{a\in\mathcal{A}}$; sampler $G_\psi$; number of paths $M$.

**Output:** Conditional-mean forecasts $\widehat{\boldsymbol{\mu}}^{\boldsymbol{a}'}(\mathbb{H}_t) \in \mathbb{R}^\tau$.

$z_{\text{init}} \leftarrow f^c_{\theta_f}(\mathbb{H}_t)$
$z_t^{(m)} \leftarrow z_{\text{init}}, m = 1, \dots, M$
**for** $r = 0, \dots, \tau - 1$ **do**
  $s \leftarrow t + r$
  $\widehat{\mu}^{\boldsymbol{a}'}_{s+1}(\mathbb{H}_t) \leftarrow M^{-1} \sum_{m=1}^{M} g_{a'_s}(z_s^{(m)})$
  **if** $r < \tau - 1$ **then**
    **for** $m = 1, \dots, M$ **do**
      $\xi_{s,m} \sim P_\xi$
      $z_{s+1}^{(m)} \leftarrow G_\psi(z_s^{(m)}, a'_s; \xi_{s,m})$
    **end for**
  **end if**
**end for**
**return** $(\widehat{\mu}^{\boldsymbol{a}'}_{t+1}(\mathbb{H}_t), \dots, \widehat{\mu}^{\boldsymbol{a}'}_{t+\tau}(\mathbb{H}_t))$

sample the full outcome distribution. Increasing $M$ reduces Monte Carlo error when the relevant moments exist, while identification and model error remain separate sources of uncertainty.

An interventional interpretation requires the longitudinal identification assumptions and representation-level mean and transition sufficiency conditions in Appendix H. One-step mean sufficiency alone is insufficient for multi-step rollout. These conditions apply to histories reached under the treatment plan; transfer further requires the relevant mechanisms and support in the target environment.

**Evaluation targets.** Simulated and semi-synthetic data provide oracle potential outcomes for evaluating intervention forecasts and treatment contrasts. On real EHR cohorts, we instead supply the recorded future treatment sequence and compare forecasts with observed outcomes in held-out environments. The resulting errors measure factual forecasting performance; direct assessment of treatment-effect error requires oracle contrasts.

### A.4 Hyperparameter and Training Summary

The default history encoder is a two-layer GRU with hidden dimension 128 and representation dimensions $(d_c, d_e) = (64, 32)$. Its per-step input follows the temporal convention above, with a zero vector or fixed start token for $a_0$. The denoiser is a three-layer MLP with hidden dimension 512. Unless stated otherwise, these architectures are shared across the benchmarks.

Stage 1 uses Adam with learning rate $3 \times 10^{-4}$, batch size 64, and 50 epochs, with $\lambda_{\text{env}} = 0.5$, $\lambda_{\text{suf}} = 0.1$, and sufficiency slack $\delta = 0$. Stage 2 uses the same learning rate, batch size 128, and 30 epochs. The contrastive temperature and weight are $\kappa_{\text{con}} = 0.2$ and $\lambda_{\text{con}} = 0.1$.

The diffusion model uses $K = 1000$ noising steps with a linear schedule from $\beta_1 = 10^{-4}$ to $\beta_K = 0.02$. Sampling uses 20 DDIM updates with $\eta = 0$ and a standard Gaussian initialization. Forecasts average $M = 100$ latent rollout paths. Appendix B specifies the sampling indices and update equations.

**Treatment encoding.** Binary-treatment datasets use $\mathcal{A} = \{0, 1\}$. Tumor Growth uses the four joint radiotherapy/chemotherapy categories defined above, except for baselines whose original interface requires two binary inputs.

**Compute.** All experiments use a single NVIDIA RTX A6000 GPU. On the real MIMIC-IV cohorts, Stage 1 takes approximately 25 minutes for $T = 168$ and 10 minutes for $T = 120$. The reported Stage 2 runtime is approximately 8 minutes under the default configuration.

# B Latent Diffusion and Counterfactual Sampling Details

This section specifies the reverse sampler $G_\psi$ used by Algorithm 2. The training objectives are defined in Appendix A.2; here we describe the forward coefficients, DDIM updates, and differentiation through the sampler.

## B.1 Forward process

With the encoder and outcome heads frozen, let $x_0 = \mathbf{z}_{t+1}^c$ denote the next state encoded from a factual history. For a variance schedule $0 < \beta_k < 1$, define $\alpha_k = 1 - \beta_k$, $\bar{\alpha}_k = \prod_{j=1}^{k} \alpha_j$, and $\bar{\alpha}_0 = 1$. The forward process has marginal distribution

$$q(x_k \mid x_0) = \mathcal{N}\big(\sqrt{\bar{\alpha}_k} x_0, (1 - \bar{\alpha}_k)\mathbf{I}\big), \qquad x_k = \sqrt{\bar{\alpha}_k} x_0 + \sqrt{1 - \bar{\alpha}_k}\epsilon. \tag{22}$$

Here $k$ is uniform on $\{1, \dots, K\}$ and $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ is independent of the factual transition. These samples define the noise-prediction loss in Eq. equation 13. The next latent supplies the denoising target, and the current latent and treatment provide the conditioning information.

## B.2 Conditional reverse sampling

Let $\mathcal{S} = (k_J, \dots, k_0)$, with $K = k_J > \cdots > k_0 = 0$, denote the reverse timestep sequence, and let $\eta \in [0, 1]$ control the DDIM sampling noise. An initialization rule $r_{\text{init}}$ maps a Gaussian draw to the initial reverse state:

$$\xi_0 \sim \mathcal{N}(0, \mathbf{I}), \qquad x_{k_J} = r_{\text{init}}(\mathbf{z}_t^c, a; \xi_0). \tag{23}$$

A standard Gaussian initialization corresponds to $r_{\text{init}}(z, a; \xi_0) = \xi_0$. The initialization, timestep sequence, and noise parameter jointly specify the sampler.

For a reverse step from $k = k_j$ to $l = k_{j-1}$, the $\epsilon$-prediction DDIM update is

$$\begin{aligned}
\widehat{\epsilon}_j &= \epsilon_\psi(x_k, k \mid \mathbf{z}_t^c, a), \\
\widehat{x}_{0,j} &= \frac{x_k - \sqrt{1 - \bar{\alpha}_k}\widehat{\epsilon}_j}{\sqrt{\bar{\alpha}_k}}, \\
\sigma_j &= \eta \sqrt{\frac{1 - \bar{\alpha}_l}{1 - \bar{\alpha}_k}} \sqrt{1 - \frac{\bar{\alpha}_k}{\bar{\alpha}_l}}, \\
x_l &= \sqrt{\bar{\alpha}_l}\widehat{x}_{0,j} + \sqrt{1 - \bar{\alpha}_l - \sigma_j^2}\widehat{\epsilon}_j + \sigma_j \zeta_j, \qquad \zeta_j \sim \mathcal{N}(0, \mathbf{I}).
\end{aligned} \tag{24}$$

The conditioning pair $(\mathbf{z}_t^c, a)$ is held fixed throughout one clinical transition. Updates are deterministic conditional on the initialization when $\eta = 0$; otherwise, the $\zeta_j$ are independent Gaussian draws. The sampler returns $G_\psi(\mathbf{z}_t^c, a; \xi) = x_{k_0}$, where $\xi$ collects all base-noise draws. If latent rescaling is used, it is included in $G_\psi$; the equations above introduce no additional rescaling.

## B.3 Differentiation through generated transitions

For the alternative-treatment sample in Eq. equation 14, $G_\psi$ is differentiated through the selected sequence of reverse updates. The contrastive gradient therefore propagates from the generated projection through the sampled state to $\psi$. The base-noise draws are independent of the parameters, and the encoded states remain detached. Together with updates to $h$ and $q$, this implements the joint Stage 2 objective in Eq. equation 17.

The shared $\mathbf{z}_t^e$ enters the projections in Eq. equation 15 during training. At inference, Algorithm 3 applies the transition sampler and outcome heads to $\mathbf{z}^c$ alone.

## B.4 Diffusion configuration

We use $K = 1000$ noising steps with the linear variance schedule of Ho et al. (2020):

$$\beta_k = 10^{-4} + \frac{k - 1}{999}(0.02 - 10^{-4}), \qquad k = 1, \dots, 1000. \tag{25}$$

The resulting terminal coefficient is $\bar{\alpha}_K \approx 4.04 \times 10^{-5}$. The reverse process starts from $x_{1000} \sim \mathcal{N}(0, \mathbf{I})$ and uses $\eta = 0$ with

$$\mathcal{S} = (1000, 950, 900, \ldots, 50, 0).$$

This sequence requires 20 denoiser evaluations; the clean index 0 is the final output and is not passed to the denoiser. The updates are deterministic conditional on the initial Gaussian draw. Independent initializations provide the $M = 100$ paths used to estimate the plan-specific mean forecasts in Algorithm 3. The Stage 1 sufficiency slack is fixed at $\delta = 0$.

# C Full Related Work

**Longitudinal ITE Estimation.** Longitudinal causal inference studies treatment effects under sequential decision-making, where time-varying confounding and history dependence complicate identification. Classical approaches such as marginal structural models and the sequential g-formula address this problem through reweighting or outcome modeling under sequential ignorability (Robins et al., 2000; Hernán & Robins, 2010). Recent deep learning methods extend ITE estimation to sequential settings by learning history representations and counterfactual outcome predictors. RMSNs combine recurrent outcome models with time-dependent inverse-propensity weighting for multi-step potential outcome forecasting (Lim, 2018). CRN learns treatment-balanced sequential representations through adversarial objectives to improve counterfactual prediction under alternative treatment paths (Bica et al., 2020b). The Time Series Deconfounder introduces latent sequential confounders inferred from multi-cause treatment structure to mitigate hidden confounding (Bica et al., 2020a). SyncTwin constructs a synthetic twin in representation space to estimate counterfactual outcomes and assess reliability in longitudinal settings (Qian et al., 2021). G-Net formulates sequential g-computation with deep outcome models for counterfactual prediction under dynamic treatment regimes (Li et al., 2020). Causal Transformer uses attention-based sequence modeling to estimate counterfactual outcomes over time (Melnychuk et al., 2022), while TE-CDE studies continuous-time counterfactual modeling with neural controlled differential equations (Seedat et al., 2022). Earlier continuous-time treatment-response models also study counterfactual reasoning with continuous-time and continuous-valued interventions using dynamical treatment-response structure (Soleimani et al., 2017).

Several recent methods are especially close to our representation-learning motivation. DCRN decomposes longitudinal representations into treatment, outcome, and confounding factors for treatment-effect inference over time (Berrevoets et al., 2021). DR-CRN also studies disentangled representation learning for sequential treatment effect estimation (Chu et al., 2022). These works show that factorized recurrent representations can improve longitudinal counterfactual prediction. DeCaf differs in the role assigned to the decomposition. Rather than decomposing factors primarily around treatment, outcome, and confounding structure, DeCaf separates an outcome-relevant component from an environment-related component under policy and environment shift. The outcome-relevant subspace $\mathbf{z}_t^c$ is used for potential-outcome prediction and treatment-conditional rollout, while the environment/nuisance subspace $\mathbf{z}_t^e$ absorbs policy-, cohort-, and measurement-related variation. This makes the decomposition directly tied to robust multi-horizon counterfactual forecasting under shift.

**Causal and Invariant Representation Learning.** A central theme in causal machine learning is to learn representations that support counterfactual generalization and transfer across environments (Schölkopf et al., 2021). In ITE estimation, representation balancing has been widely used to reduce distribution mismatch between treated and control groups. Johansson et al. (2016) proposed learning representations motivated by domain adaptation, and Shalit et al. (2017) provided generalization bounds and practical algorithms based on integral probability metrics. SITE further preserves local neighborhood structure in addition to global balancing, improving individualized estimation (Yao et al., 2018). Another related line of work models latent confounding using deep generative latent-variable inference, such as CEVAE (Louizos et al., 2017), and the deconfounder framework formalizes this idea through multiple causes (Wang & Blei, 2019). Factorization-based methods also address noisy or missing covariates in causal learning (Kallus et al., 2018).

Beyond treatment-effect estimation, invariant prediction and domain generalization provide related principles for robustness under distribution shift. Invariant causal prediction searches for relationships that remain stable across environments (Peters et al., 2016), while invariant risk minimization encourages predictors that are simultaneously optimal across environments (Arjovsky et al., 2019). Group distributionally robust optimization improves worst-group performance under known group

shifts (Sagawa et al., 2019). Domain-adversarial training learns representations that reduce domain information through gradient reversal (Ganin et al., 2016), and domain separation networks explicitly split shared and private features across domains (Bousmalis et al., 2016). Weakly supervised disentanglement further shows that meaningful factorization usually requires supervision, inductive bias, or structural constraints (Locatello et al., 2020). These ideas motivate the use of environment labels and structural objectives in DECAF.

The main distinction is that DECAF ties the decomposition directly to longitudinal causal prediction rather than using invariance as a generic domain-generalization objective (Peters et al., 2016; Arjovsky et al., 2019; Sagawa et al., 2019; Schölkopf et al., 2021). The outcome heads and transition model operate on $\mathbf{z}_t^c$, while $\mathbf{z}_t^e$ is trained to retain environment-related variation, following the broader idea of domain separation but adapting it to longitudinal treatment-effect estimation (Ganin et al., 2016; Bousmalis et al., 2016; Locatello et al., 2020). Moreover, DECAF does not rely on invariance alone. Because an invariant representation can discard outcome-relevant information when domain and outcome factors are entangled, we combine environment separation with an outcome-sufficiency regularizer that penalizes residual predictive gain from the raw history beyond $\mathbf{z}_t^c$ (Arjovsky et al., 2019; Locatello et al., 2020; Schölkopf et al., 2021). This design aims to reduce shortcut learning, where models rely on unstable but predictive artifacts that may fail under shift (Geirhos et al., 2020).

**Policy and Environment Shift in Clinical Time Series.** Longitudinal EHR data are especially vulnerable to policy and environment shift because treatment policies, measurement frequency, documentation practices, and patient populations can vary across hospitals, cohorts, care units, and time periods (Johnson et al., 2023; Liu et al., 2020; 2023). A model trained in one environment may therefore learn correlations that reflect local practice rather than stable treatment response, a challenge closely related to invariant prediction, shortcut learning, and domain generalization (Peters et al., 2016; Arjovsky et al., 2019; Geirhos et al., 2020; Schölkopf et al., 2021). This issue is especially important for treatment-effect estimation because the learned representation is used not only for one-step prediction but also for counterfactual rollout under alternative treatment plans (Robins et al., 2000; Hernán & Robins, 2010; Lim, 2018; Bica et al., 2020b; Melnychuk et al., 2022). When the rollout state entangles treatment-response information with policy- or cohort-specific artifacts, an intervention on treatment can inadvertently propagate environment-specific variation over future steps. DECAF addresses this setting by using environment labels during training to separate environment-related variation from the latent component used for treatment-effect prediction and rollout, drawing on domain-adversarial and domain-separation objectives but applying them to longitudinal causal prediction (Ganin et al., 2016; Bousmalis et al., 2016; Locatello et al., 2020).

**Counterfactual Generation.** Generative modeling provides a mechanism for synthesizing counterfactual samples and is increasingly used for causal estimation and robustness. In cross-sectional settings, CEVAE supports counterfactual outcome sampling through latent confounder modeling (Louizos et al., 2017), while GANITE imputes counterfactuals through adversarial generation (Yoon et al., 2018). Normalizing-flow-based approaches are appealing because of invertibility and density learning: diffeomorphic counterfactual constructions study counterfactual search in flow-induced coordinate systems (Dombrowski et al., 2023), and causal normalizing flows provide a way to implement interventions and answer counterfactual queries within learned structural models (Javaloy et al., 2023). PO-Flow learns densities of potential outcomes and counterfactuals through continuous normalizing flows (Wu et al., 2025a), while FCCL uses normalizing flows and contrastive learning to align factual and counterfactual representations for robust ITE estimation (Zhang et al., 2025).

For temporal settings, counterfactual generation has been explored from multiple perspectives. Counterfactual generative models for time-varying treatments directly model potential outcome trajectories under alternative treatment paths (Wu et al., 2024). CDVAE combines dynamic variational autoencoding with propensity-based weighting for counterfactual regression in longitudinal data (Bouchattaoui et al., 2023). DoFlow models multivariate time series with a causal DAG and continuous normalizing flows to support interventional and counterfactual trajectory prediction (Wu et al., 2025b). Structured state-space models have also been explored for counterfactual outcome prediction in sequential settings (Purohit, 2023). Separately, work on counterfactual explanations for forecasting formulates counterfactual time-series perturbations under validity and manifold constraints (Wang et al., 2023).

Diffusion models have become effective probabilistic generators for time-series forecasting (Rasul et al., 2021), building on general denoising and score-based diffusion models (Ho et al., 2020; Song et al., 2020). Recent causal diffusion models further apply diffusion to counterfactual outcome

Table 6: Construction of the in-distribution (ID) and out-of-distribution (OOD) splits. OOD environments are excluded from training, validation, and model selection.

| Dataset | Training and validation environments | ID test | OOD test |
|---|---|---|---|
| Tumor Growth | Three source policies with $\gamma \in \{0.5, 1.0, 1.5\}$; patient-disjoint training and validation sets | Held-out patients under the same source policies, $\gamma \in \{0.5, 1.0, 1.5\}$ | Test-only policies: OOD-1, $\gamma = 2.0$; OOD-2, $\gamma = 3.0$; OOD-3, $\gamma = 4.0$ |
| Semi-MIMIC-III | SCM environments 0–3: 1,620 training and 200 validation patients | 180 held-out source patients, obtained by an environment-stratified 10% holdout from the initial 1,800-patient training pool | SCM environments 4–7: 2,000 test patients |
| MIMIC-IV ventilation | MICU, MICU/SICU, and SICU: 2,144 training and 267 validation stays | 271 held-out stays from the same three care units | All 686 CVICU and 732 TSICU stays (1,418 total) |

distributions in longitudinal data (Alinezhad et al., 2026). TIDAL studies temporal causal diffusion for longitudinal medical imaging outcomes (Cuturrufo et al.), showing that diffusion-based causal modeling is also relevant beyond scalar clinical time series.

DECAF uses this generative perspective differently. Rather than using a generator to impute missing potential outcomes directly or to generate the full outcome trajectory, DECAF uses generation as a latent transition operator. The generator learns $p_\psi(\mathbf{z}^c_{t+1} \mid \mathbf{z}^c_t, \mathrm{a}_t)$ and is applied only to the subspace used by the outcome heads. During counterfactual rollout, $\mathbf{z}^e_t$ is kept fixed so that the intervention changes the outcome-relevant state without directly perturbing environment-, cohort-, or measurement-related variation. This separates the benefit of generative temporal modeling from the benefit of operating in a decomposed causal subspace.

# D Additional Experimental Details

## D.1 Definition and Construction of the OOD Splits

We use *environment OOD* to mean that the histories in the test set are drawn from environments that are absent from model fitting, validation, hyperparameter selection, and normalization-statistic estimation. This is distinct from a counterfactual query: imposing a new treatment sequence on a history from a training environment creates an off-policy counterfactual query, but does not by itself create an environment shift. Table 6 summarizes the split used for each benchmark.

**Tumor Growth: treatment-policy shift.** We define each environment by its treatment-assignment coefficient $\gamma_e$. The three source environments use $\gamma_e = 0.5$, $1.0$, and $1.5$, corresponding to weak, mild, and moderate dependence of treatment assignment on patient history. Within each source environment, patients are partitioned into training, validation, and ID-test sets before temporal examples are constructed. ID testing therefore evaluates unseen patients under policies represented during training.

The held-out environments use $\gamma_e = 2.0$, $3.0$, and $4.0$, denoted OOD-1, OOD-2, and OOD-3. These coefficients represent mild, strong, and severe extrapolation beyond the source range, respectively; the labels refer to policy distance rather than an assumed ordering of prediction error. All OOD patients and policies are excluded from fitting, validation, hyperparameter selection, and normalization-statistic estimation. Patient partitions are mutually disjoint, and all windows and counterfactual branches remain in the partition of their originating patient.

The tumor-response equations, patient-response parameter distribution, initial-state distribution, drug-exposure dynamics, process noise, and observation protocol are identical across environments. Observation noise does not define separate environments. Only the history-dependent assignment policy changes. For each method and training seed, one model is trained jointly on the three source

environments, selected using source validation, and applied without adaptation to every ID and OOD policy. The source policy defines the environment label used in training; neither that label nor $\gamma_e$ is supplied to the outcome heads or transition sampler at inference.

Counterfactual query construction is separate from this environment split. The policy coefficient generates the factual history before the forecast origin. Future actions are then imposed by the queried treatment plan, rather than sampled from the observational policy. Single-sliding and random treatment sequences are evaluated on both ID and OOD histories using the same query-construction procedure across policies and methods. Oracle simulator outcomes permit $\mathrm{nRMSE}_{cf}(\tau)$ evaluation for $\tau \in \{1, \ldots, 5\}$ in both settings.

**Semi-MIMIC-III.** Our Semi-MIMIC-III evaluation uses 4,000 patient trajectories of length 30, organized into eight environments indexed by $e \in \{0, \ldots, 7\}$, with each environment corresponding to a distinct SCM parameterization. Environments 0–3 constitute the source environments, and environments 4–7 are reserved for OOD evaluation. The source environments contain 2,000 patients. Of these, 200 are assigned to validation, and the remaining 1,800 form the initial training pool. A fixed, environment-stratified 10% holdout from this pool provides 180 ID-test patients, leaving 1,620 patients for model fitting. The 2,000 patients in environments 4–7 form the OOD-test set. For the reported ID/OOD comparison, the same models trained on the 1,620-patient training split are evaluated on both test sets, without refitting or adaptation.

The benchmark provides oracle one-step potential outcomes for both treatment arms. We evaluate next-step treatment-effect estimation using $\sqrt{\mathrm{PEHE}}$ and ATE error. At each prediction origin, the model uses only the factual history available before the current treatment decision. Dataset construction and evaluation details are provided in Appendix D.4.

**MIMIC-IV: held-out care units.** Care unit defines the real-data environment. After cohort construction, the 4,100 stays comprise 1,035 MICU, 701 MICU/SICU, 946 SICU, 686 CVICU, and 732 TSICU stays. Only the first three care units are used as source environments. Within these source units, patient-level splitting gives 2,144 training, 267 validation, and 271 ID-test stays. CVICU and TSICU are never used during training or model selection; all 686 CVICU and 732 TSICU stays form the 1,418-stay OOD test set. All preprocessing and normalization statistics are estimated from the 2,144-stay training split and then applied unchanged to the validation, ID-test, and OOD-test sets. Splitting is performed at the stay level before forecast windows are constructed. For the patient-level analysis in Table 7, we deterministically take the earliest eligible 24-hour window from each OOD stay, so no patient receives additional weight through multiple windows. Illustrative gallery cases are selected only from this fixed window set and are reported separately from the aggregate result. This 24-hour patient-level summary is reported separately from the horizon-specific forecasting results in the main text.

Because MIMIC-IV does not provide counterfactual outcomes, its OOD experiment evaluates factual $SpO_2$ forecasting conditional on the recorded future ventilation sequence. It measures generalization to held-out care units, not counterfactual treatment-effect identification. Moreover, the observed ID and held-out-unit errors are similar; we therefore describe this as a *held-out care-unit evaluation*, without claiming that the realized shift is necessarily large.

This appendix provides dataset construction details, evaluation protocols, baseline configurations, metric definitions, and additional ablations. The main paper reports representative results, while this appendix gives the full experimental context needed for reproduction and interpretation.

### D.2 Tumor Growth Simulator

**Simulator and treatments.** We follow the Tumor Growth simulator and evaluation protocol used in prior longitudinal counterfactual prediction work (Geng et al., 2017; Lim, 2018; Bica et al., 2020b; Melnychuk et al., 2022). The simulator models a one-dimensional tumor volume $Y_t$ over discrete time steps with two binary treatments: radiotherapy $A_t^r$, which has an immediate effect, and chemotherapy $A_t^c$, which has a delayed effect.

Table 7: Detailed patient-level factual forecasting in held-out MIMIC-IV care units. We report mean patient error with patient-bootstrap 95% confidence intervals in brackets. Each of the 1,418 stays contributes one 24-hour window, and errors are computed only at genuinely observed future $SpO_2$ hours. CVICU and TSICU contain 686 and 732 stays; the stable-treatment and treatment-transition groups contain 1,171 and 247 stays, respectively. Lower is better, and bold denotes the lowest mean error in each column.

| Method | All | CVICU | TSICU | Stable | Transition |
|---|---|---|---|---|---|
| *Panel A: MAE ($SpO_2$ percentage points)* | | | | | |
| **DECAF** | **1.673 [1.634, 1.714]** | **1.770 [1.708, 1.830]** | **1.582 [1.530, 1.634]** | **1.642 [1.596, 1.689]** | **1.821 [1.727, 1.912]** |
| CRN | 1.693 [1.651, 1.734] | 1.799 [1.740, 1.860] | 1.594 [1.541, 1.649] | 1.656 [1.610, 1.702] | 1.869 [1.785, 1.960] |
| Unsplit-DDPM | 1.756 [1.717, 1.795] | 1.817 [1.761, 1.876] | 1.699 [1.651, 1.749] | 1.731 [1.689, 1.776] | 1.872 [1.790, 1.959] |
| Persistence | 1.962 [1.877, 2.054] | 2.180 [2.044, 2.356] | 1.758 [1.667, 1.862] | 1.865 [1.774, 1.972] | 2.421 [2.220, 2.637] |
| *Panel B: nRMSE* | | | | | |
| **DECAF** | **0.683 [0.667, 0.701]** | **0.729 [0.704, 0.757]** | **0.640 [0.620, 0.662]** | **0.670 [0.652, 0.689]** | **0.746 [0.710, 0.781]** |
| CRN | 0.687 [0.672, 0.705] | 0.734 [0.710, 0.759] | 0.644 [0.621, 0.668] | 0.673 [0.654, 0.692] | 0.757 [0.723, 0.791] |
| Unsplit-DDPM | 0.708 [0.692, 0.724] | 0.741 [0.718, 0.765] | 0.678 [0.659, 0.697] | 0.698 [0.680, 0.716] | 0.757 [0.723, 0.791] |
| Persistence | 0.811 [0.781, 0.844] | 0.888 [0.839, 0.942] | 0.738 [0.704, 0.773] | 0.774 [0.740, 0.810] | 0.983 [0.917, 1.060] |

Table 8: **Benchmark roles and evaluation.** Oracle counterfactuals are available for Tumor Growth and semi-synthetic MIMIC-III, but not for real MIMIC-IV.

| **Dataset / Ground truth** | **Evaluation objective** | **Protocol / Metrics** |
|---|---|---|
| **Tumor Growth**<br>Synthetic; oracle trajectories | Multi-step counterfactual accuracy under held-out treatment policies | Source $\gamma \in \{0.5, 1.0, 1.5\}$;<br>OOD $\gamma \in \{2.0, 3.0, 4.0\}$;<br>$\tau \in \{1, \ldots, 5\}$<br>$\mathrm{nRMSE}_{cf}(\tau) \downarrow$ |
| **Semi-MIMIC-III**<br>Semi-synthetic; oracle effects | Treatment-effect estimation under environment shift | ID and held-out environments; $\tau = 1$<br>$\sqrt{\mathrm{PEHE}} \downarrow$; $\mathrm{ATEErr} \downarrow$ (OOD) |
| **MIMIC-IV**<br>Real; factual outcomes only | Factual forecasting under care-unit shift | Held-out CVICU and TSICU<br>Horizon-specific<br>$\mathrm{nRMSE}_{\mathrm{forecast}}(\tau) \downarrow$;<br>24-hour patient-level MAE and nRMSE $\downarrow$ (Table 7) |

**Outcome dynamics.** Tumor volume evolves according to a shared pharmacokinetic and pharmacodynamic model:

$$Y_{t+1} = \left(1 + \rho \log\left(\frac{K_{\mathrm{tum}}}{Y_t}\right) - \beta_c C_t - (\alpha_r d_t + \beta_r d_t^2) + \epsilon_t\right) Y_t, \qquad \epsilon_t \sim \mathcal{N}(0, 0.01^2). \quad (26)$$

Here, $K_{\mathrm{tum}}$ is the tumor carrying capacity, distinct from the number of diffusion steps. The radiotherapy dose $d_t$ has an immediate effect, while chemotherapy exposure $C_t$ includes the current dose and residual exposure from previous doses. The Gaussian disturbance enters the volume update as $Y_t \epsilon_t$. Patient-specific response parameters, including $\rho$, $\beta_c$, $\alpha_r$, and $\beta_r$, follow the same population sampling law in every environment; treatment-response heterogeneity follows the standard simulator parameterization (Melnychuk et al., 2022).

All environments share this response equation, the parameter distribution, initial-state distribution, chemotherapy decay rule, dose levels, process-noise law, observation protocol, and boundary and stopping rules. Thus, environments may contain different patients but do not differ in biological response mechanisms. Mechanism invariance is understood conditional on the full simulator state, including residual drug exposure and patient-specific parameters. It does not require the observational conditional distribution $P(Y_{t+1} \mid Y_t, A_t)$ to be identical across environments, since current tumor volume alone omits these state variables.

**Treatment assignment and confounding.** The only environment-dependent component is the treatment-assignment policy. Let $D_s = (6Y_s/\pi)^{1/3}$ denote tumor diameter and let

$$\overline{D}_{15,t} = \frac{1}{w_t} \sum_{s=\max(1,t-14)}^{t} D_s, \qquad w_t = \min(15, t), \tag{27}$$

be its recent-history average available before treatment $A_t$. This indexing follows the convention that $Y_t$ is observed before $A_t$ is assigned and that $A_t$ affects $Y_{t+1}$. In environment $e$, the two treatments are conditionally independent Bernoulli draws with probability

$$\begin{aligned} p_e(\mathbb{H}_t) &= \sigma\left[\frac{\gamma_e}{D_{\max}}\left(\overline{D}_{15,t} - \frac{D_{\max}}{2}\right)\right], \\ A_t^c, A_t^r \mid \mathbb{H}_t, e &\overset{\text{ind}}{\sim} \text{Bernoulli}(p_e(\mathbb{H}_t)). \end{aligned} \tag{28}$$

Here $\sigma$ is the sigmoid function, and the same $D_{\max}$ and policy midpoint are used in all environments. The coefficient $\gamma_e$ controls how strongly treatment assignment responds to recent tumor burden; it does not change treatment efficacy. Larger coefficients increase assignment sensitivity, without necessarily increasing the marginal treatment rate.

**Evaluation settings.** We train jointly on source policies $\Gamma_{\text{src}} = \{0.5, 1.0, 1.5\}$ and evaluate at horizons $\tau \in \{1, 2, 3, 4, 5\}$. ID evaluation uses held-out patients from each source policy. OOD-1, OOD-2, and OOD-3 use the test-only coefficients 2.0, 3.0, and 4.0, respectively. All model fitting, preprocessing, and selection use source data only, and the same fitted model is evaluated across all policies without adaptation. For multi-step queries, we retain the single sliding treatment interventions and random treatment trajectories used in the benchmark protocol (Melnychuk et al., 2022). Simulator branches preserve each patient's response parameters and residual chemotherapy exposure at the forecast origin; only the future treatment sequence is intervened on.

This design evaluates transfer across treatment policies while holding biological mechanisms fixed. Different policies can nevertheless induce different distributions of patient histories and clinical states. Environment regularization is therefore paired with outcome sufficiency to reduce reliance on policy-specific associations while retaining information needed for treatment-response prediction. The design does not assume that exact marginal environment invariance of the predictive state is always attainable.

**Treatment encoding.** RMSN receives the two binary treatments $(A_t^r, A_t^c)$ as separate inputs, following its original interface (Lim, 2018). For the remaining deep baselines, the two binary treatments are encoded as a single four-category treatment variable corresponding to $\{(0,0), (1,0), (0,1), (1,1)\}$, consistent with prior comparisons (Melnychuk et al., 2022).

### D.3 Policy-wise Tumor Growth Results

Table 9 reports horizon-specific errors under a representative source policy, $\gamma = 1.0$, and the three held-out policies, $\gamma \in \{2.0, 3.0, 4.0\}$. The $\gamma = 1.0$ results use held-out patients and provide a policy-specific ID evaluation. For each method and seed, the same model trained jointly on $\Gamma_{\text{src}} = \{0.5, 1.0, 1.5\}$ is evaluated across these policies without adaptation. The three held-out policies correspond to OOD-1, OOD-2, and OOD-3, respectively.

### D.4 Semi-synthetic MIMIC-III

**Benchmark construction.** We use a semi-synthetic MIMIC-III setup adapted from the DSW framework (Liu et al., 2020), together with an environment-based evaluation protocol. Clinical covariates and treatment histories are extracted from an adult sepsis cohort in MIMIC-III, whereas outcomes are generated by an SCM. The observed covariates include time-varying clinical measurements and static demographic variables. The binary treatment indicates vasopressor administration. This construction retains observed clinical covariate and treatment patterns while providing oracle potential outcomes for treatment-effect evaluation.

Our implemented dataset contains 4,000 patient trajectories, each of length 30, across eight environments indexed by $e \in \{0, \ldots, 7\}$. Each environment corresponds to a distinct SCM parameterization $\theta_e$. The eight-environment partition described below is the evaluation protocol used in this work.

Table 9: **Tumor-growth counterfactual forecasting under treatment-policy shift.** All models are trained jointly on source policies $\gamma \in \{0.5, 1.0, 1.5\}$. We report $\mathrm{nRMSE}_{cf}(\tau) \downarrow$ for $\tau \in \{1, \ldots, 5\}$ on held-out patients under the representative ID policy $\gamma = 1.0$ and the OOD policies $\gamma = 2.0, 3.0, 4.0$. DeCaf uses four joint treatment arms; its entries are mean $\pm$ sample standard deviation over seeds 10, 101, and 1010. Lower is better.

| $\gamma$ | Method | $\tau = 1$ | $\tau = 2$ | $\tau = 3$ | $\tau = 4$ | $\tau = 5$ |
|---|---|---|---|---|---|---|
| 1 | MSM | 2.274±0.376 | 2.787±0.451 | 3.087±0.499 | 3.222±0.519 | 3.216±0.501 |
| | RMSN | 1.446±0.225 | 1.524±0.253 | 1.583±0.259 | 1.612±0.243 | 1.611±0.197 |
| | CRN | 1.300±0.127 | 1.371±0.173 | 1.431±0.196 | 1.477±0.206 | 1.520±0.205 |
| | DR-CRN | 1.693±0.121 | 1.732±0.158 | 1.761±0.179 | 1.766±0.194 | 1.746±0.177 |
| | G-Net | 1.655±0.197 | 1.949±0.307 | 2.162±0.387 | 2.353±0.451 | 2.551±0.510 |
| | CT | 1.593±0.094 | 1.709±0.096 | 1.790±0.113 | 1.828±0.125 | 1.805±0.129 |
| | G-Transformer | 2.235±0.263 | 2.699±0.349 | 3.173±0.351 | 3.362±0.489 | 3.437±0.380 |
| | CTD-NKO | 1.308±0.042 | 1.358±0.073 | 1.416±0.109 | 1.475±0.132 | 1.538±0.130 |
| | CDM | 1.203±0.013 | 1.237±0.046 | 1.412±0.204 | 1.457±0.176 | 1.578±0.175 |
| | **DeCaf (ours)** | **0.930±0.114** | **1.068±0.159** | **1.231±0.376** | **1.393±0.537** | **1.515±0.624** |
| 2 | MSM | 2.794±0.494 | 3.299±0.577 | 3.590±0.630 | 3.709±0.646 | 3.671±0.612 |
| | RMSN | 1.697±0.365 | 1.824±0.364 | 1.915±0.343 | 1.967±0.321 | 1.965±0.290 |
| | CRN | 1.491±0.284 | 1.646±0.337 | 1.746±0.353 | 1.803±0.342 | 1.826±0.305 |
| | DR-CRN | 1.716±0.281 | 1.917±0.421 | 1.967±0.445 | 1.982±0.449 | 1.984±0.423 |
| | G-Net | 1.670±0.285 | 1.972±0.446 | 2.115±0.536 | 2.181±0.579 | 2.205±0.599 |
| | CT | 1.565±0.191 | 1.746±0.263 | 1.838±0.284 | 1.951±0.292 | 2.089±0.353 |
| | G-Transformer | 2.725±0.458 | 3.363±0.482 | 3.852±0.710 | 4.110±0.751 | 4.061±0.671 |
| | CTD-NKO | 1.402±0.053 | 1.480±0.091 | 1.526±0.126 | 1.576±0.151 | **1.643±0.177** |
| | CDM | 1.302±0.024 | 1.342±0.093 | 1.534±0.024 | 1.693±0.023 | 1.689±0.342 |
| | **DeCaf (ours)** | **0.972±0.137** | **1.104±0.135** | **1.293±0.238** | **1.498±0.350** | 1.662±0.457 |
| 3 | MSM | 3.149±0.491 | 3.543±0.583 | 3.780±0.630 | 3.892±0.652 | 3.857±0.627 |
| | RMSN | 1.675±0.332 | 1.820±0.403 | 2.007±0.549 | 2.121±0.597 | 2.129±0.561 |
| | CRN | 1.920±0.427 | 2.237±0.536 | 2.393±0.569 | 2.447±0.568 | 2.405±0.525 |
| | DR-CRN | 1.680±0.223 | 1.958±0.369 | 2.200±0.420 | 2.383±0.424 | 2.514±0.385 |
| | G-Net | 1.940±0.120 | 2.459±0.161 | 2.870±0.314 | 3.219±0.566 | 3.510±0.892 |
| | CT | 2.194±0.203 | 2.536±0.320 | 2.691±0.357 | 2.745±0.418 | 2.654±0.434 |
| | G-Transformer | 3.428±0.574 | 4.263±0.633 | 4.828±0.750 | 5.063±0.691 | 4.970±0.641 |
| | CTD-NKO | 1.561±0.136 | 1.738±0.198 | 1.885±0.242 | 1.995±0.264 | 2.075±0.264 |
| | CDM | 1.521±0.858 | 1.631±0.012 | 1.929±0.258 | 2.942±0.271 | 2.578±0.824 |
| | **DeCaf (ours)** | **1.431±0.297** | **1.275±0.166** | **1.439±0.263** | **1.770±0.359** | **2.067±0.468** |
| 4 | MSM | 3.243±0.575 | 3.497±0.664 | 3.744±0.737 | 3.884±0.770 | 3.844±0.754 |
| | RMSN | 2.597±0.647 | 2.912±0.737 | 3.121±0.840 | 3.181±0.897 | 3.094±0.867 |
| | CRN | 2.203±0.630 | 2.681±0.789 | 2.925±0.904 | 3.023±0.973 | 2.991±0.982 |
| | DR-CRN | 2.592±0.297 | 3.334±0.449 | 3.642±0.529 | 3.668±0.535 | 3.517±0.504 |
| | G-Net | 2.246±0.277 | 2.920±0.442 | 3.373±0.502 | 3.700±0.456 | 3.914±0.419 |
| | CT | 2.092±0.195 | 2.335±0.228 | 2.542±0.211 | 2.724±0.192 | 2.849±0.187 |
| | G-Transformer | 4.431±1.032 | 5.388±1.079 | 6.006±1.107 | 6.257±0.958 | 6.150±0.903 |
| | CTD-NKO | **1.963±0.280** | 2.296±0.440 | 2.536±0.544 | 2.673±0.604 | 2.707±0.604 |
| | CDM | 2.632±0.035 | 2.834±0.025 | 2.902±0.295 | 3.421±0.632 | 3.678±0.244 |
| | **DeCaf (ours)** | 2.150±0.333 | **2.129±0.232** | **2.137±0.300** | **2.196±0.424** | **2.277±0.497** |

**Potential outcomes and prediction target.** At each evaluated decision time $t$, the benchmark provides oracle next-step potential outcomes $y_{t+1}^{(i)}(0)$ and $y_{t+1}^{(i)}(1)$ for patient $i$. These correspond to assigning treatment 0 or 1 at time $t$ while leaving the patient's factual history unchanged. The oracle treatment-effect contrast and its model estimate are

$$\begin{aligned} \Delta_t^{(i)} &= y_{t+1}^{(i)}(1) - y_{t+1}^{(i)}(0), \\ \widehat{\Delta}_t^{(i)} &= \widehat{y}_{t+1}^{(i)}(1) - \widehat{y}_{t+1}^{(i)}(0). \end{aligned} \tag{29}$$

The model constructs its pre-treatment representation using only the factual history available before the treatment decision at time $t$. Measurements after that decision are not used to construct the

representation. The oracle potential outcomes serve as evaluation targets; they are not inputs to the history encoder.

**Environment partition.** We divide the eight environments into source and target sets:

$$\mathcal{E}_{\mathrm{src}} = \{0, 1, 2, 3\}, \qquad \mathcal{E}_{\mathrm{tgt}} = \{4, 5, 6, 7\}. \tag{30}$$

The source environments contain 2,000 patients in total and are used for model development and in-distribution evaluation. The target environments contain 2,000 additional patients and are reserved exclusively for out-of-distribution evaluation.

**In-distribution evaluation (ID/IID).** The ID test set evaluates generalization to unseen patients from the same source environments used for training. This is the source-distribution test setting, also referred to here as IID testing. The terminology concerns patient trajectories and does not require independence between successive observations within a trajectory.

Within the source environments, 200 patients are assigned to validation and the remaining 1,800 form an initial training pool. We reserve a fixed, environment-stratified 10% holdout from this pool for ID testing. The final source-domain partition therefore consists of 1,620 training patients, 200 validation patients, and 180 ID-test patients. The ID-test patients are excluded from model fitting, but their generating environments are represented in the training set.

**Out-of-distribution evaluation (OOD).** The OOD-test set consists of all 2,000 patients from environments 4–7. Neither these patients nor their environment-specific SCM parameterizations are used during model development. OOD evaluation therefore assesses transfer to held-out generating environments. The distinction between ID and OOD is the generating environment: ID testing holds out patients within the source environments, whereas OOD testing holds out both patients and SCM parameterizations. OOD membership is determined by this partition and does not imply that its prediction error must exceed the ID error.

**Model fitting and selection.** All models are fitted using only the 1,620 training patients. Early stopping and hyperparameter selection use only the 200 source-validation patients. Data-dependent preprocessing, including normalization, is fitted on the training split and applied unchanged to the validation and test splits.

For each training seed, the same fitted model is evaluated on both the 180-patient ID-test set and the 2,000-patient OOD-test set, without refitting or target-environment adaptation. The training, validation, ID-test, and OOD-test splits are mutually disjoint at the patient level. All methods use the same patient partitions.

**Evaluation metrics.** We evaluate one-step treatment-effect estimation at forecasting horizon $\tau = 1$. We report ID and OOD $\sqrt{\mathrm{PEHE}}$ and OOD ATE error, using the oracle potential outcomes and the metric definitions given in the evaluation appendix.

For a descriptive comparison between the two test settings, the rooted-PEHE gap is defined as

$$\Delta_{\mathrm{ID/OOD}} = \sqrt{\mathrm{PEHE}}_{\mathrm{OOD}} - \sqrt{\mathrm{PEHE}}_{\mathrm{ID}}. \tag{31}$$

Both terms are computed using the same fitted model. We report absolute ID and OOD errors alongside this gap, because the source and target environments may differ in intrinsic prediction difficulty. The gap alone is therefore not used to establish robustness to environment shift.

### D.5 Real MIMIC-IV Mechanical Ventilation Cohort

We construct a real-world MIMIC-IV mechanical ventilation cohort to evaluate factual forecasting on held-out ICU care units. We report horizon-specific forecasting results and a complementary 24-hour patient-level analysis in Table 7. Since oracle counterfactual outcomes are unavailable in real EHR data, these evaluations assess factual forecasting performance rather than counterfactual treatment-effect accuracy. Following prior work on ICU respiratory-support forecasting, we build temporally aligned patient trajectories from MIMIC-IV v2.2 using ICU stay information, demographics, admissions, charted vital signs, and ventilation procedure records (Johnson et al., 2023; Moukheiber et al., 2024a;b). The resulting cohort contains hourly clinical covariates, invasive mechanical ventilation indicators, and future $SpO_2$ outcomes.

**Cohort construction.** We construct an adult ICU cohort using the first ICU stay for each patient and include stays from MICU, MICU/SICU, SICU, CVICU, and TSICU. To support long-horizon forecasting, we retain stays with at least 7 days of hourly observations and sufficient vital-sign coverage, including at least 144 valid hours of vital measurements and 96 valid hours of $SpO_2$. After filtering, the cohort contains $N = 4{,}100$ ICU stays: 1,035 from MICU, 701 from MICU/SICU, 946 from SICU, 686 from CVICU, and 732 from TSICU.

**Environment split.** We use ICU care unit as the environment label. MICU, MICU/SICU, and SICU are used as training environments, while CVICU and TSICU are held out as OOD environments. The 2,682 source-environment stays are split into 2,144 training, 267 validation, and 271 ID-test stays (approximately 80/10/10). The 686 CVICU and 732 TSICU stays are pooled into the OOD-test set, yielding 1,418 OOD stays. Models are selected using the validation split from the training environments and evaluated on both ID-test and OOD-test splits.

**Treatment and outcome.** The treatment is invasive mechanical ventilation. We mark an hour as treated, $a_t = 1$, when an invasive ventilation episode is active during that hourly interval. The forecasting target is the per-hour mean $SpO_2$. The horizon-specific forecasting protocol uses future horizons up to $\tau_{\max} = 40$ hours; the patient-level analysis in Table 7 uses a fixed 24-hour evaluation window. In our implementation, ventilation episodes are extracted from `procedureevents.csv` using itemid 225792, and $SpO_2$ measurements are extracted from `chartevents.csv` using itemid 220277.

**Covariates and temporal alignment.** We use routinely measured ICU variables that are clinically relevant for respiratory-support forecasting, including heart rate, respiratory rate, $SpO_2$, non-invasive and arterial blood pressure, temperature, $FiO_2$, and PEEP. Measurements are clipped to clinically reasonable ranges and aggregated into hourly bins relative to ICU admission, producing fixed-length sequences of $T = 168$ hours. Missing values are imputed using last observation carried forward within each stay. For reproducibility, the corresponding MIMIC-IV itemids are: heart rate 220045, respiratory rate 220210, $SpO_2$ 220277, non-invasive blood pressure 220179/220180/220181, arterial blood pressure 220050/220051/220052, temperature 223762, $FiO_2$ 223835, and PEEP 220339.

**Autoregressive $SpO_2$ history and leakage prevention.** $SpO_2$ is included both as a historical covariate and as the forecasting target. This does not introduce information leakage because the model input at decision time $t$ contains only measurements observed up to time $t$, while the target is a future $SpO_2$ value at $t + \tau$. Thus, past $SpO_2$ provides an autoregressive clinical signal, but future $SpO_2$ measurements are never included in the history $\mathbb{H}_t$ used for prediction. All continuous covariates and outcomes are standardized using statistics computed only from the training split of the training environments, and the same statistics are applied unchanged to validation, ID-test, and OOD-test stays.

**Patient-level evaluation on held-out care units.** For Table 7, each of the 1,418 OOD stays contributes exactly one window: the earliest eligible 24-hour window specified in Section D.1. The care-unit groups contain 686 CVICU and 732 TSICU stays; the stable-treatment and treatment-transition groups contain 1,171 and 247 stays, respectively. Forecasts are conditioned on the recorded future ventilation sequence. Errors are computed only at genuinely observed future $SpO_2$ hours; imputed future target values are not scored in this analysis. We summarize MAE in $SpO_2$ percentage points and normalized RMSE as mean patient errors, with patient-bootstrap 95% confidence intervals. These intervals quantify patient-sampling uncertainty and are distinct from standard deviations across training seeds.

# E Additional Ablations and Diagnostics

We examine the contribution of the training objectives, transition family, and representation dimensions. All Tumor Growth variants use the same three source policies, held-out test patients, forecast horizons $\tau \in \{1, \ldots, 5\}$, and three training seeds as the main experiment. The loss and transition-family tables summarize the four reported test policies, $\gamma \in \{1, 2, 3, 4\}$; these averages combine the representative ID policy with the three OOD policies.

### E.1 Loss Ablations on Tumor Growth

We ablate each major component of the DeCaf objective. Table 10 compares the full model with variants that omit environment separation, outcome sufficiency, or counterfactual-pair regularization. Removing $\mathcal{L}_{\text{env}}$ or $\mathcal{L}_{\text{suf}}$ changes Stage 1, whereas removing $\mathcal{L}_{\text{con}}$ changes Stage 2. For fixed Stage 1 encoder and outcome-head parameters, all transition variants have the same one-step prediction $g_a(f^c_{\theta_f}(\mathbb{H}_t))$. Their transition-dependent differences are therefore evaluated at horizons greater than one.

Table 10: Loss ablation on Tumor Growth. Entries are mean $\text{nRMSE}_{cf}(\tau) \downarrow$ over the four reported test policies, $\gamma \in \{1, 2, 3, 4\}$, and three training seeds. The Avg. column averages the displayed horizons $\{1, 3, 5\}$. This table reports means only. Lower is better.

| Method | $\tau = 1$ | $\tau = 3$ | $\tau = 5$ | Avg. |
|---|---|---|---|---|
| Full DeCaf | 1.371 | 1.525 | 1.880 | 1.592 |
| w/o $\mathcal{L}_{\text{env}}$ | 1.990 | 1.870 | 1.951 | 1.937 |
| w/o $\mathcal{L}_{\text{suf}}$ | 1.993 | 1.671 | 1.912 | 1.859 |
| w/o $\mathcal{L}_{\text{con}}$ | 1.371 | 1.923 | 1.997 | 1.764 |

### E.2 Transition Family and Latent-Space Ablations

We compare multiple transition families under the same learned representation and also evaluate whether the transition model should operate in the outcome-relevant subspace rather than an entangled or nuisance representation. Table 11 separates the effect of the transition model from the effect of the latent space. The $\mathbf{z}^e$ rollout is used as a diagnostic rather than a competitive baseline: it tests whether the nuisance branch alone contains stable treatment-response information. The key controlled comparison is between rollout in $\mathbf{z}^c$ and rollout in the unsplit latent $\mathbf{z}$ under the same transition family.

Table 11: **Transition family and latent-space ablation on Tumor Growth.** Entries are mean $\text{nRMSE}_{cf}(\tau) \downarrow$ over the four reported test policies, $\gamma \in \{1, 2, 3, 4\}$, and three training seeds. This table reports means only. Lower is better.

| Space | Transition | $\tau = 1$ | $\tau = 2$ | $\tau = 3$ | $\tau = 4$ | $\tau = 5$ |
|---|---|---|---|---|---|---|
| $\mathbf{z}^c$ | DECAF | 1.371 | 1.394 | 1.525 | 1.714 | 1.880 |
| $\mathbf{z}^c$ | Gaussian | 1.371 | 2.185 | 2.639 | 3.045 | 3.439 |
| $\mathbf{z}^c$ | CVAE | 1.371 | 1.504 | 1.765 | 1.985 | 2.200 |
| $\mathbf{z}^c$ | Flow | 1.371 | 1.787 | 2.190 | 2.534 | 2.839 |
| Unsplit $\mathbf{z}$ | DDPM | 1.475 | 1.483 | 1.618 | 1.912 | 2.175 |
| $\mathbf{z}^e$ | DDPM | 34.764 | 28.151 | 26.470 | 25.963 | 25.773 |

### E.3 Frozen-Probe Diagnostics on Real MIMIC-IV

We use frozen-representation probes to check whether the rollout subspace $\mathbf{z}^c_t$ preserves outcome information while reducing treatment-policy predictability. Table 12 shows that the main configuration keeps $\mathbf{z}^c_t$ close to the full-history outcome-prediction ceiling, retaining $97.9\%$ of the history-level $R^2$, while reducing treatment-policy predictability from $\mathbf{z}^c_t$ by $9.6$ percentage points. This supports using $\mathbf{z}^c_t$ as the outcome-relevant rollout subspace under policy shift.

### E.4 Dimension Sensitivity

We evaluate how the dimensions of the outcome-relevant and environment-related branches affect counterfactual forecasting. The default configuration used in the main comparisons is $(d_c, d_e) =$

Table 12: **Frozen-probe diagnostics on MIMIC-IV.** The main configuration preserves outcome information in $\mathbf{z}_t^c$ while reducing treatment-policy predictability. The table reports point summaries; seed-level standard deviations are not provided.

| Role of $\mathbf{z}_t^c$ | Diagnostic | Main result | Takeaway |
|---|---|---|---|
| Outcome-relevant subspace | $R^2(y\|\mathbf{z}_t^c)$ vs. $R^2(y\|\mathbb{H}_t)$ | 0.826 vs. 0.844 | retains 97.9% of history ceiling |
| Less policy-dependent rollout state | $\mathrm{Acc}(\mathrm{a}_t\|\mathbf{z}_t^c)$ | $95.4\% \rightarrow 85.8\%$ | 9.6 pp lower policy predictability |

$(64, 32)$. Table 13 shows that larger latent spaces do not consistently improve accuracy. The configurations $(32, 16)$, $(64, 32)$, and $(128, 64)$ vary both branches and therefore assess joint capacity rather than the isolated effect of $d_c$. In the comparisons that hold $d_c = 64$ fixed, increasing $d_e$ from 16 to 64 increases the reported mean error in every displayed setting. The default configuration attains the lowest reported mean at $\gamma = 4$ and $\tau = 5$, but is not uniformly best. These results indicate sensitivity to the allocation of capacity between the branches, rather than a general advantage for larger representations.

Table 13: Dimension sensitivity on Tumor Growth. We report $\mathrm{nRMSE}_{cf}(\tau) \downarrow$ for the representative ID policy $\gamma = 1.0$ and OOD-3 policy $\gamma = 4.0$. All configurations use the same three source policies and four joint treatment arms, without counterfactual supervision or oracle model selection. Only $(d_c, d_e)$ varies; all other hyperparameters are fixed. Entries are mean $\pm$ sample standard deviation over three training seeds. Lower is better.

| $d_c$ | $d_e$ | $\gamma$ | $\tau = 1$ | $\tau = 5$ |
|---|---|---|---|---|
| 32 | 16 | 1 | **0.798 ± 0.102** | 1.519 ± 0.026 |
| 64 | 32 | 1 | 0.930 ± 0.114 | 1.515 ± 0.624 |
| 128 | 64 | 1 | 1.053 ± 0.190 | 1.678 ± 0.275 |
| 64 | 16 | 1 | 0.924 ± 0.173 | **1.415 ± 0.117** |
| 64 | 64 | 1 | 0.959 ± 0.206 | 2.088 ± 0.728 |
| 32 | 16 | 4 | 1.740 ± 0.703 | 2.427 ± 0.854 |
| 64 | 32 | 4 | 2.150 ± 0.333 | **2.277 ± 0.497** |
| 128 | 64 | 4 | 2.396 ± 0.657 | 4.540 ± 2.213 |
| 64 | 16 | 4 | **1.590 ± 0.485** | 2.360 ± 0.175 |
| 64 | 64 | 4 | 2.729 ± 0.236 | 3.824 ± 0.670 |

# F Baseline Details

We compare against classical longitudinal causal estimators and deep sequential counterfactual models.

**MSM-IPW (Robins et al., 2000).** Marginal structural models with inverse probability weighting estimate treatment effects by reweighting observed trajectories according to treatment propensities.

**Sequential g-formula (Robins et al., 2000).** The sequential g-formula models the outcome process and rolls outcomes forward under specified treatment regimes.

**RMSN (Lim, 2018).** Recurrent Marginal Structural Networks combine recurrent outcome models with time-dependent inverse probability weighting for multi-step counterfactual prediction.

**CRN (Bica et al., 2020b).** Counterfactual Recurrent Networks learn treatment-conditioned sequential representations with adversarial balancing to reduce treated and untreated distribution mismatch.

**DR-CRN (Chu et al., 2022).** DR-CRN learns disentangled recurrent representations for sequential treatment-effect estimation. We include DR-CRN as the closest disentanglement-based recurrent baseline to DeCaf. Unlike DR-CRN, which focuses on disentangled sequential treatment-effect

estimation, DECAF explicitly separates outcome-relevant and environment-related latent components and performs treatment-conditional rollout only in the outcome-relevant subspace.

**G-Net (Li et al., 2020).** G-Net is a deep sequential g-computation model that learns the data generating process and simulates outcomes under dynamic treatment regimes.

**Causal Transformer (Melnychuk et al., 2022).** Causal Transformer uses attention-based sequence modeling for longitudinal counterfactual outcome prediction.

**TARNet (Shalit et al., 2017).** TARNet combines a shared representation with treatment-specific outcome heads. It provides an outcome-regression comparison for one-step treatment effects.

**G-Transformer (Xiong et al., 2024).** G-Transformer uses a Transformer-based model of time-varying covariates and Monte Carlo g-computation to forecast outcomes under treatment regimes.

**CDM (Alinezhad et al., 2026).** The Causal Diffusion Model uses conditional denoising to model counterfactual outcome distributions in longitudinal data. This baseline contrasts outcome diffusion with DECAF's diffusion transitions in a decomposed latent state. We use the abbreviation CDM consistently in the result tables.

**TE-CDE (Seedat et al., 2022).** Treatment Effect Neural Controlled Differential Equations model continuous-time latent trajectories and use adversarial training to address time-dependent confounding. We include TE-CDE in the real MIMIC-IV comparison.

**Persistence.** This forecasting reference repeats the last observed outcome at every future horizon and has no learned parameters.

**CTD-NKO (Wang et al., 2025).** CTD-NKO is included as a recent neural baseline for longitudinal counterfactual forecasting. We use the public/default configuration when available and match the same data splits and evaluation horizons.

**Baseline hyperparameters.** Baselines use the shared default configuration adapted from Melnychuk et al. (2022). For the real MIMIC-IV cohort, methods within each comparison use the same held-out care-unit split and evaluation protocol. Per-policy results are summarized over three training seeds. For ID and OOD macro-averages, policies are averaged within each seed before computing the mean and sample standard deviation, as defined in Appendix G. The loss and transition-family ablation tables report means over their four displayed test policies and three seeds. Table 7 instead reports mean patient errors with patient-bootstrap 95% confidence intervals for its fixed 24-hour windows. Baseline hyperparameters follow their default configurations adapted from prior longitudinal counterfactual forecasting work, without extensive dataset-specific tuning.

## G METRIC DEFINITIONS

**Counterfactual endpoint error.** For the policy-shift evaluation, let $\mathcal{Q}_{e,\tau,c}$ be the set of eligible patient–origin–plan queries in environment $e$, at horizon $\tau$, for query family $c$. The two families are single-sliding interventions and random treatment trajectories, and their scores are reported separately. For query $j$, write $i_j$ for the patient, $t_j$ for the forecast origin, and $\mathbf{a}^{(j)} \in \mathcal{A}^\tau$ for the prescribed future plan. The endpoint score is

$$\mathrm{nRMSE}_{cf}(e,\tau,c) = \frac{1}{\sigma_{Y,\mathrm{src}}} \sqrt{\frac{1}{|\mathcal{Q}_{e,\tau,c}|} \sum_{j \in \mathcal{Q}_{e,\tau,c}} \left( Y_{t_j+\tau}^{(i_j),\mathbf{a}^{(j)}} - \widehat{\mu}_{t_j+\tau}^{(i_j),\mathbf{a}^{(j)}} \right)^2}. \tag{32}$$

The denominator $\sigma_{Y,\mathrm{src}}$ is estimated from factual source-training outcomes and held fixed across policies and horizons. The numerator and denominator use the same outcome scale. Each eligible query receives equal weight; equal patient weighting additionally requires the same query count per patient. All methods use the same query sets and eligibility rules.

This metric measures the error at time $t+\tau$, rather than a cumulative error over the first $\tau$ forecast steps. A horizon-$\tau$ prediction uses $\tau-1$ latent transitions. A treatment plan contains $\tau$ joint actions; a constant-treatment plan $(a,\dots,a)$ is a special case. Equation equation 32 pools squared errors across the eligible queries before taking the square root.

**Rooted PEHE.** The semi-synthetic MIMIC-III experiment evaluates binary treatment effects at $\tau=1$. Let $\mathcal{I}$ index the evaluated patient–decision-time pairs and write $\Delta_j = Y^{(i_j)}_{t_j+1}(1) - Y^{(i_j)}_{t_j+1}(0)$ and $\widehat{\Delta}_j = \widehat{Y}^{(i_j)}_{t_j+1}(1) - \widehat{Y}^{(i_j)}_{t_j+1}(0)$. The rooted effect-prediction error is

$$\sqrt{\mathrm{PEHE}} = \sqrt{\frac{1}{|\mathcal{I}|}\sum_{j\in\mathcal{I}}(\Delta_j - \widehat{\Delta}_j)^2}. \tag{33}$$

**ATE error.** Using the same evaluation pairs and weights, the average-effect error is

$$\mathrm{ATEErr} = \left| \frac{1}{|\mathcal{I}|}\sum_{j\in\mathcal{I}}\Delta_j - \frac{1}{|\mathcal{I}|}\sum_{j\in\mathcal{I}}\widehat{\Delta}_j \right|. \tag{34}$$

**Factual forecasting error.** For real-world cohorts where oracle counterfactual outcomes are unavailable, we report factual forecasting error:

$$\mathrm{nRMSE}_{\mathrm{forecast}}(\tau) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y^i_{t+\tau} - \widehat{Y}^i_{t+\tau}\right)^2}}{\sigma_{Y,\mathrm{src}}}. \tag{35}$$

Here $N$ counts eligible factual patient–origin pairs at horizon $\tau$, and $\sigma_{Y,\mathrm{src}}$ is estimated from source-training outcomes. The same normalization is used for ID and OOD evaluation.

**Patient-level factual forecasting summary.** The horizon-specific error above is distinct from the patient-level summary in Table 7. For that table, each patient contributes the earliest eligible 24-hour window, and each patient's error is evaluated only over genuinely observed future $SpO_2$ hours within that window. MAE is expressed in $SpO_2$ percentage points; nRMSE uses the outcome standard deviation estimated from the training split. We report the mean of the patient-level errors in the full OOD cohort and in each reported subgroup, so each patient has equal weight within the relevant group. Bracketed intervals are patient-bootstrap 95% confidence intervals, rather than standard deviations over random seeds. Consequently, these window-averaged patient-level summaries should not be interpreted as the error at a single forecasting horizon.

**Aggregation across policies and training seeds.** Let $R_{r,e}(\tau,c)$ denote a score for training seed $r$, environment $e$, horizon $\tau$, and query family $c$. For a set of policies $G$, define the within-seed macro-average

$$R_{r,G}(\tau,c) = \frac{1}{|G|}\sum_{e\in G} R_{r,e}(\tau,c). \tag{36}$$

ID and OOD summaries are computed separately with $G=\Gamma_{\mathrm{src}}$ and $G=\Gamma_{\mathrm{ood}}$, respectively. For the three seeds $\{10,101,1010\}$, the reporting convention is

$$\overline{R}_G = \frac{1}{3}\sum_r R_{r,G}, \qquad s_G = \sqrt{\frac{1}{2}\sum_r (R_{r,G} - \overline{R}_G)^2}, \tag{37}$$

where $(\tau,c)$ is suppressed. Fixed-policy results use the same calculation with $G$ containing one policy. Thus policy averaging precedes the computation of training-seed variation. A pooled standard deviation across policy–seed combinations has a different interpretation because it includes between-policy differences.

Table 9 reports the representative ID policy $\gamma=1.0$ separately from the three OOD policies. Tables 2, 10, and 11 report means over $G=\{1,2,3,4\}$ and the three seeds. These four-policy summaries describe the selected evaluation policies and are distinct from the ID and OOD macro-averages defined above. These tables report means only; pooled variation across policies and seeds is not reported as the seed-level standard deviation of a macro-average. Patient-bootstrap confidence intervals in the MIMIC-IV analysis quantify patient-sampling uncertainty separately from training-seed variation. Scores are rounded to three decimal places.

**ID/OOD gap.** A gap compares the same metric for the same fitted model in the two domains. Consistent with the one-step semi-synthetic evaluation, the rooted-PEHE gap is

$$\Delta_{\mathrm{ID/OOD}} = \sqrt{\mathrm{PEHE}}_{\mathrm{OOD}}(\tau = 1) - \sqrt{\mathrm{PEHE}}_{\mathrm{ID}}(\tau = 1). \tag{38}$$

For Tumor Growth, the analogous gap compares an OOD-policy score or OOD macro-average with the ID macro-average at the same horizon and query family. Differences are calculated within each seed before summarizing across seeds. Gaps are reported with both absolute errors because the source and target populations may differ in prediction difficulty. A negative gap does not by itself establish robustness.

# H Identification assumptions and theoretical proofs

## H.1 Standard longitudinal identification assumptions

We use the following standard assumptions for step-wise longitudinal causal identification (Robins, 1986; Robins et al., 2000; Hernán & Robins, 2010). All conditional statements below are understood almost surely on the relevant support. Histories take values in standard Borel spaces, the representations are finite-dimensional, and the relevant outcomes have finite second moments.

**Assumption 2** (Standard longitudinal identification)**.** For each time $t$ and treatment level $a \in \mathcal{A}$, assume: *Consistency:* if $\mathrm{a}_t = a$, then the observed outcome equals the corresponding potential outcome, $\mathrm{y}_{t+1} = \mathrm{y}^a_{t+1}$. *Sequential ignorability:* $\mathrm{y}^a_{t+1} \perp\!\!\!\perp \mathrm{a}_t \mid \mathbb{H}_t$. *Positivity:* $0 < P(\mathrm{a}_t = a \mid \mathbb{H}_t) < 1$ for all histories with positive probability. *No interference/SUTVA:* treatments have no multiple versions, and one unit's treatment does not affect another unit's outcome.

Under Assumption 2, the conditional mean potential outcome is identifiable as

$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbb{H}_t].$$

Past treatments are included in $\mathbb{H}_t$, so the conditioning history already encodes the observed treatment path before time $t$. These conditions identify one-step potential-outcome means. Multi-step identification additionally requires consistency, sequential exchangeability, and positivity for the future covariate and outcome process at each decision time along the evaluated treatment plan; one-step outcome exchangeability alone does not identify the entire transition process.

## H.2 Representation level decomposition and identification

We now state the representation-level condition used in the main text. The condition specifies the target role of the learned representation: $\mathbf{z}^c_t$ should retain the information needed for conditional mean potential-outcome prediction, while $\mathbf{z}^e_t$ should contain variation that is not needed once $\mathbf{z}^c_t$ is known. This is a representation-level statement and does not require recovering the underlying clinical, biological, or physical state.

**Assumption 3** (Representation level decomposition)**.** For each $t$, there exist random variables $(\mathbf{z}^c_t, \mathbf{z}^e_t)$ such that, almost surely and for all $a \in \mathcal{A}$:

1. *Measurability.* $\mathbf{z}^c_t$ and $\mathbf{z}^e_t$ are measurable with respect to $\sigma(\mathbb{H}_t)$.
2. *Representation-level mean sufficiency.* There exists a measurable function $m_{t,a}$ such that

$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbb{H}_t] = m_{t,a}(\mathbf{z}^c_t).$$

Assumption 3 is stated at the level of conditional mean potential outcomes, which is the target estimated by the treatment-specific outcome heads. It is stronger than what can be verified from finite samples, but it gives a clean population-level condition under which the outcome heads can operate on $\mathbf{z}^c_t$ rather than on the full history $\mathbb{H}_t$.

**Proposition 2** (Existence of a measurable encoder)**.** *Under Assumption 3(1), there exists a measurable mapping $f$ such that*

$$f(\mathbb{H}_t) = (\mathbf{z}^c_t, \mathbf{z}^e_t) \quad \textit{a.s.}$$

*Proof.* Assumption 3(1) states that $\mathbf{z}_t^c$ and $\mathbf{z}_t^e$ are measurable with respect to $\sigma(\mathbb{H}_t)$. By the Doob-Dynkin factorization lemma, there exist measurable functions $f_t^c$ and $f_t^e$ such that

$$\mathbf{z}_t^c = f_t^c(\mathbb{H}_t), \qquad \mathbf{z}_t^e = f_t^e(\mathbb{H}_t)$$

almost surely. Defining $f(h) = (f_t^c(h), f_t^e(h))$ gives the claim. □

**Corollary 3** (Outcome sufficiency of $\mathbf{z}_t^c$)**.** *Under Assumption 3, for each $a \in \mathcal{A}$,*

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c, \mathbf{z}_t^e] = \mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c] = m_{t,a}(\mathbf{z}_t^c) \quad \textit{a.s.}$$

*Thus, nuisance mean irrelevance and the absence of residual conditional-mean information in the history follow from measurability and mean sufficiency. Moreover, if the conditional mean potential-outcome mechanism is invariant across environments given $\mathbf{z}_t^c$, namely*

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c, \mathrm{e} = e] = m_{t,a}(\mathbf{z}_t^c)$$

*on the relevant support for each $e \in \mathcal{E}$, then $m_{t,a}$ is a common conditional-mean predictor across those environments.*

*Proof.* Fix $a \in \mathcal{A}$. By Assumption 3(1), both $\sigma(\mathbf{z}_t^c)$ and $\sigma(\mathbf{z}_t^c, \mathbf{z}_t^e)$ are contained in $\sigma(\mathbb{H}_t)$. Using the tower property and Assumption 3(2),

$$\begin{aligned}\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c, \mathbf{z}_t^e] &= \mathbb{E}\big[\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbb{H}_t] \mid \mathbf{z}_t^c, \mathbf{z}_t^e\big] \\ &= \mathbb{E}[m_{t,a}(\mathbf{z}_t^c) \mid \mathbf{z}_t^c, \mathbf{z}_t^e] = m_{t,a}(\mathbf{z}_t^c).\end{aligned}$$

The same argument conditioning only on $\mathbf{z}_t^c$ gives

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c] = m_{t,a}(\mathbf{z}_t^c).$$

Together with Assumption 3(2), these identities prove the stated equalities. The environment-invariance statement follows from the additional equality of conditional means across environments; it is not implied by marginal invariance of $\mathbf{z}_t^c$ alone. □

**Theorem 4** (Representation level identification)**.** *Suppose Assumption 2 holds. Let $\mathbf{z}_t^c = f_t^c(\mathbb{H}_t)$ be a measurable representation. If for each $a \in \mathcal{A}$ there exists a function $m_a$ such that*

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbb{H}_t] = m_a(\mathbf{z}_t^c),$$

*then*

$$m_a(\mathbf{z}_t^c) = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}_t^c].$$

*Equivalently,*

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}_t^c].$$

*Proof.* Fix $a \in \mathcal{A}$. By consistency and sequential ignorability given $\mathbb{H}_t$,

$$\mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbb{H}_t].$$

Taking conditional expectations with respect to $(\mathrm{a}_t = a, \mathbf{z}_t^c)$ gives

$$\mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}_t^c] = \mathbb{E}\big[\mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbb{H}_t] \mid \mathrm{a}_t = a, \mathbf{z}_t^c\big].$$

Using the previous equality and the assumed mean sufficiency,

$$\mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}_t^c] = \mathbb{E}\big[\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbb{H}_t] \mid \mathrm{a}_t = a, \mathbf{z}_t^c\big] = \mathbb{E}[m_a(\mathbf{z}_t^c) \mid \mathrm{a}_t = a, \mathbf{z}_t^c] = m_a(\mathbf{z}_t^c).$$

Positivity ensures that the treatment-specific conditional mean is defined on almost every history of interest and that the equality is not restricted to histories observed only under another treatment. This proves the result. □

**Theorem 5** (Environment invariant Bayes predictor)**.** *Fix $a \in \mathcal{A}$. Assume that, for almost every $z$ and every $e \in \mathcal{E}$,*

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c = z, \mathrm{e} = e] = \mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c = z].$$

*Then the Bayes regressor under squared loss is*

$$g_a^\star(z) = \mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{z}_t^c = z],$$

*and this regressor is identical across environments on the relevant support. The statement does not identify predictions outside that support.*

*Proof.* For squared loss, the Bayes regressor is the conditional mean. For any predictor $g$, the conditional risk satisfies

$$\mathbb{E}\big[(\mathrm{y}^a_{t+1} - g(\mathbf{z}^c_t))^2 \mid \mathbf{z}^c_t, \mathrm{e} = e\big] = \mathrm{Var}(\mathrm{y}^a_{t+1} \mid \mathbf{z}^c_t, \mathrm{e} = e) + \big(\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbf{z}^c_t, \mathrm{e} = e] - g(\mathbf{z}^c_t)\big)^2 .$$

The variance term does not depend on $g$. By the assumed conditional invariance,

$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbf{z}^c_t, \mathrm{e} = e] = \mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbf{z}^c_t]$$

for all environments. Thus the minimizer is the same in every environment and equals

$$g^\star_a(\mathbf{z}^c_t) = \mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbf{z}^c_t].$$

□

**Proposition 6** (Population interpretation of the Stage 1 objective)**.** *Consider the population version of the Stage 1 objective with universal function classes for the treatment-specific outcome heads, the environment adversary, and the sufficiency probes. Let*

$$(\mathbf{z}^c_t, \mathbf{z}^e_t) = f(\mathbb{H}_t), \qquad Z = \mathbf{z}^c_t,$$

*and define the factual conditional means*

$$\mu_H = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t, \mathbb{H}_t], \qquad \mu_Z = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t, Z].$$

*The optimal squared prediction risks are*

$$R^\star_Z = \inf_{\{h_a\}} \mathbb{E}\big[(\mathrm{y}_{t+1} - h_{\mathrm{a}_t}(Z))^2\big] ,$$

$$R^\star_H = \inf_{\{\tilde{h}_a\}} \mathbb{E}\Big[(\mathrm{y}_{t+1} - \tilde{h}_{\mathrm{a}_t}(\mathbb{H}_t))^2\Big] ,$$

*where the infima are over measurable predictors. Then*

$$R^\star_Z - R^\star_H = \mathbb{E}[(\mu_H - \mu_Z)^2].$$

*Consequently, equality of the risks implies $\mu_H = \mu_Z$ almost surely under the observed joint distribution. Under treatment positivity, this yields, for each $a \in \mathcal{A}$,*

$$\mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}^c_t]$$

*for almost every supported history. If Assumption 2 holds, the same equality identifies*

$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbb{H}_t] = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbf{z}^c_t].$$

*Separately, for finite environment labels and an unrestricted environment classifier trained by cross-entropy, the optimal risk using $Z$ equals the risk of the marginal environment predictor if and only if $Z \perp\!\!\!\perp \mathrm{e}$.*

*Proof.* Under squared loss and universal function classes, the Bayes predictors are the corresponding conditional means $\mu_H$ and $\mu_Z$. Since $Z$ is a measurable function of history, $\sigma(\mathrm{a}_t, Z)$ is contained in $\sigma(\mathrm{a}_t, \mathbb{H}_t)$. Expand the residual as

$$\mathrm{y}_{t+1} - \mu_Z = (\mathrm{y}_{t+1} - \mu_H) + (\mu_H - \mu_Z).$$

The conditional expectation of $\mathrm{y}_{t+1} - \mu_H$ given $(\mathrm{a}_t, \mathbb{H}_t)$ is zero, while $\mu_H - \mu_Z$ is measurable with respect to this conditioning sigma-field. The cross term therefore vanishes in expectation, giving

$$\mathbb{E}[(\mathrm{y}_{t+1} - \mu_Z)^2] = \mathbb{E}[(\mathrm{y}_{t+1} - \mu_H)^2] + \mathbb{E}[(\mu_H - \mu_Z)^2].$$

This proves the risk identity. If the risks are equal, the final nonnegative term has expectation zero and hence $\mu_H = \mu_Z$ almost surely.

To obtain a treatment-specific statement, let

$$d_a(\mathbb{H}_t) = \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, \mathbb{H}_t] - \mathbb{E}[\mathrm{y}_{t+1} \mid \mathrm{a}_t = a, Z], \qquad \pi_a(\mathbb{H}_t) = P(\mathrm{a}_t = a \mid \mathbb{H}_t).$$

The risk identity can be written as

$$R^\star_Z - R^\star_H = \sum_{a \in \mathcal{A}} \mathbb{E}[\pi_a(\mathbb{H}_t) d_a(\mathbb{H}_t)^2].$$

When this quantity is zero and $\pi_a(\mathbb{H}_t) > 0$ almost surely, $d_a(\mathbb{H}_t) = 0$ for almost every supported history. Consistency and conditional exchangeability in Assumption 2 then give the potential-outcome mean identity.

For the environment statement, the optimal cross-entropy classifier predicts $P(\mathrm{e} \mid Z)$ and attains conditional entropy $H(\mathrm{e} \mid Z)$. The marginal predictor attains $H(\mathrm{e})$. Their difference is the mutual information

$$H(\mathrm{e}) - H(\mathrm{e} \mid Z) = I(\mathrm{e}; Z),$$

which is zero exactly when $Z \perp\!\!\!\perp \mathrm{e}$. □

This proposition gives a population-level interpretation of the empirical Stage 1 losses. The outcome and sufficiency terms encourage conditional mean sufficiency of $\mathbf{z}_t^c$ for treatment-specific prediction, while the adversarial environment term encourages marginal removal of environment-predictive information from $\mathbf{z}_t^c$. The environment head on $\mathbf{z}_t^e$ encourages the complementary branch to retain environment-predictive variation. Universal function classes and population-optimal probes are essential to this interpretation; a finite empirical probe gap need not equal the population risk difference. Even a small population difference controls an average weighted by observed treatment frequencies rather than uniform error on rare treatment-history combinations. The argument does not establish that environment invariance and outcome sufficiency are simultaneously attainable.

**Derivation of the latent g-formula.** Consider a fixed future treatment plan $\bar{a} = (a_t, \ldots, a_{t+\tau-1})$. Assume consistency and sequential exchangeability for the future history and outcome process at each decision time, together with positivity on every history reached by the plan. In addition to terminal outcome mean sufficiency, assume transition sufficiency: for each relevant $s$ and $a$, the interventional next-latent law depends on the current history only through $\mathbf{z}_s^c$,

$$\mathcal{L}(\mathbf{z}_{s+1}^{c,a} \mid \mathbb{H}_s) = K_{s,a}(\cdot \mid \mathbf{z}_s^c), \qquad \mathbf{z}_{s+1}^{c,a} = f_{s+1}^c(\mathbb{H}_{s+1}^a).$$

These conditions are required along the entire plan, including histories reached after earlier interventions. Then the conditional mean at horizon $\tau$ has the latent g-formula

$$\mathbb{E}[\mathrm{y}_{t+\tau}^{\bar{a}} \mid \mathbb{H}_t] = \int m_{t+\tau-1,a_{t+\tau-1}}(z_{t+\tau-1}) \prod_{s=t}^{t+\tau-2} K_{s,a_s}(\mathrm{d}z_{s+1} \mid z_s), \qquad z_t = \mathbf{z}_t^c, \tag{39}$$

where the integral is over the intermediate latent states and $m_{s,a}$ denotes the time-specific conditional mean in Assumption 3. For $\tau = 1$, the kernel product is empty and the right-hand side is $m_{t,a_t}(\mathbf{z}_t^c)$.

*Proof.* For a bounded measurable function $\phi$, consistency and sequential exchangeability imply

$$\begin{aligned}\int \phi(z') K_{s,a}(\mathrm{d}z' \mid \mathbf{z}_s^c) &= \mathbb{E}[\phi(\mathbf{z}_{s+1}^{c,a}) \mid \mathbb{H}_s] \\ &= \mathbb{E}[\phi(\mathbf{z}_{s+1}^c) \mid \mathrm{a}_s = a, \mathbb{H}_s].\end{aligned}$$

Taking conditional expectations given $(\mathrm{a}_s = a, \mathbf{z}_s^c)$ shows that $K_{s,a}$ equals the observed treatment-specific latent transition kernel on the relevant support. At the last decision time, outcome mean sufficiency expresses the conditional mean under the final treatment as $m_{t+\tau-1,a_{t+\tau-1}}(\mathbf{z}_{t+\tau-1}^c)$. Successive conditioning backward through the interventional history process replaces each next-latent conditional distribution by its kernel $K_{s,a_s}$. Iterating this step from $s = t + \tau - 2$ down to $s = t$ gives Eq. equation 39. Finite moments ensure that the outcome expectations are well-defined. □

Transition sufficiency is an additional condition beyond one-step outcome mean sufficiency. The derivation concerns the population kernels and does not assert that a fitted diffusion sampler equals those kernels.

### H.3 Linear Gaussian recovery of the causal subspace

We next give a tractable Gaussian setting showing how invariance and outcome sufficiency can recover the treatment-outcome-relevant part of the latent state. The Gaussian formulation makes the argument transparent because the Bayes predictors are linear and the effect of environment shifts can be read from the conditional moments. The same intuition extends to broader settings in which environment changes induce identifiable variation in nuisance directions and the outcome mechanism depends on a lower-dimensional stable component.

**Assumption 4** (Linear Gaussian decomposition)**.** Let $\mathbf{s}_t \in \mathbb{R}^{d_c}$ be a latent causal state and $\mathbf{u}_t \in \mathbb{R}^{d_e}$ be a latent nuisance state. Assume:

1. *Linear mixing.* The observable history embedding $\mathbf{h}_t \in \mathbb{R}^d$ satisfies
$$\mathbf{h}_t = \mathbf{A}\mathbf{s}_t + \mathbf{B}\mathbf{u}_t,$$
where $\mathbf{A} \in \mathbb{R}^{d\times d_c}$ and $\mathbf{B} \in \mathbb{R}^{d\times d_e}$ have full column rank, the concatenated matrix $[\mathbf{A}\ \mathbf{B}]$ has full column rank, and $d \geq d_c + d_e$.

2. *Outcome depends on the treatment-outcome state.* For each $a \in \mathcal{A}$,
$$\mathrm{y}^a_{t+1} = \beta_a^\top \mathbf{s}_t + \eta_t,$$
where $\eta_t$ is mean-zero noise independent of $(\mathbf{s}_t, \mathbf{u}_t, \mathrm{e})$ and satisfies $\mathbb{E}[\eta_t^2] < \infty$.

3. *Non-degenerate Gaussian state.* The state has distribution
$$\mathbf{s}_t \sim \mathcal{N}(0, \Sigma_S), \qquad \Sigma_S \succ 0,$$
and is independent of the environment label e.

4. *Environment shifts the nuisance mean.* $\mathbf{s}_t \perp\!\!\!\perp \mathrm{e}$, and conditional on $\mathrm{e} = e$,
$$\mathbf{u}_t \sim \mathcal{N}(\mu_e, \Sigma_U),$$
with covariance $\Sigma_U$ independent of $e$.

5. *Rich environments.* For some reference environment $e_0$, the set
$$\{\mu_e - \mu_{e_0} : e \in \mathcal{E}\}$$
spans $\mathbb{R}^{d_e}$. Every environment in this statement has positive probability.

6. *Predictive subspace.* Let
$$\mathcal{S}_y = \mathrm{span}\{\beta_a : a \in \mathcal{A}\}$$
be the treatment-outcome predictive subspace of $\mathbf{s}_t$. Positive definiteness of $\Sigma_S$ ensures nonzero variance along every nonzero direction in this subspace.

The Gaussian-state condition is an explicit additional distributional assumption. Centering avoids an intercept in the recovery statement; a nonzero state mean can be handled by centering or an affine formulation. No independence between $\mathbf{s}_t$ and $\mathbf{u}_t$ is required for the result.

**Theorem 7** (Linear Gaussian recovery)**.** *Under Assumption 4, consider a linear representation*
$$\mathbf{z}^c_t = \mathbf{W}^\top \mathbf{h}_t, \qquad \mathbf{W} \in \mathbb{R}^{d\times d_c},$$
*with* $\mathbf{W}$ *full column rank. Suppose* $\mathbf{z}^c_t$ *satisfies:*

1. Environment invariance: *the distribution of* $\mathbf{z}^c_t \mid \mathrm{e} = e$ *is the same for all* $e \in \mathcal{E}$.

2. Outcome sufficiency: *for each* $a \in \mathcal{A}$*, predicting* $\mathrm{y}^a_{t+1}$ *from* $\mathbf{z}^c_t$ *attains the optimal squared prediction risk obtainable from the true latent state* $\mathbf{s}_t$*:*
$$\inf_g \mathbb{E}[(\mathrm{y}^a_{t+1} - g(\mathbf{z}^c_t))^2] = \inf_f \mathbb{E}[(\mathrm{y}^a_{t+1} - f(\mathbf{s}_t))^2],$$
*where the infima are over measurable predictors.*

*Then* $\mathbf{z}^c_t$ *removes the nuisance component and preserves the treatment-outcome predictive subspace. In particular,*
$$\mathbf{W}^\top \mathbf{B} = 0,$$
*and for each* $a \in \mathcal{A}$ *there exists a vector* $\ell_a$ *such that*
$$\mathbb{E}[\mathrm{y}^a_{t+1} \mid \mathbf{s}_t] = \beta_a^\top \mathbf{s}_t = \ell_a^\top \mathbf{z}^c_t \quad \textit{a.s.}$$
*If, in addition, the predictive subspace spans the full latent state,* $\mathcal{S}_y = \mathbb{R}^{d_c}$*, then there exists an invertible matrix* $R \in \mathbb{R}^{d_c\times d_c}$ *such that*
$$\mathbf{z}^c_t = R^\top \mathbf{s}_t \quad \textit{a.s.}$$

*Proof.* Using the linear mixing model,
$$\mathbf{z}^c_t = \mathbf{W}^\top \mathbf{h}_t = \mathbf{W}^\top \mathbf{A}\mathbf{s}_t + \mathbf{W}^\top \mathbf{B}\mathbf{u}_t.$$

**Step 1: invariance removes the nuisance component.** Conditional on $\mathrm{e} = e$, the nuisance variable has mean $\mu_e$. Since $\mathbf{s}_t$ is independent of e, the environment-dependent mean term in $\mathbf{z}_t^c$ is $\mathbf{W}^\top \mathbf{B}\mu_e$. Environment invariance requires the conditional distribution of $\mathbf{z}_t^c$ to be the same across environments; in particular, its conditional mean must be the same. Thus, for all $e \in \mathcal{E}$,

$$\mathbf{W}^\top \mathbf{B}(\mu_e - \mu_{e_0}) = 0.$$

By the rich environment condition, the vectors $\{\mu_e - \mu_{e_0} : e \in \mathcal{E}\}$ span $\mathbb{R}^{d_e}$. Therefore,

$$\mathbf{W}^\top \mathbf{B} = 0.$$

Thus,

$$\mathbf{z}_t^c = \mathbf{W}^\top \mathbf{A}\mathbf{s}_t \quad \text{a.s.}$$

**Step 2: outcome sufficiency preserves the treatment-outcome predictive component.** Let

$$M = \mathbf{W}^\top \mathbf{A}.$$

After Step 1, we have

$$\mathbf{z}_t^c = M\mathbf{s}_t \quad \text{a.s.}$$

Fix $a \in \mathcal{A}$ and write $V_a = \beta_a^\top \mathbf{s}_t$. Because the outcome noise is independent of the state and has mean zero, the optimal squared prediction risk using $\mathbf{s}_t$ is

$$R_s^\star = \mathbb{E}[\eta_t^2].$$

The optimal predictor using $M\mathbf{s}_t$ is $\mathbb{E}[V_a \mid M\mathbf{s}_t]$, and its risk satisfies

$$R_z^\star = \mathbb{E}[\eta_t^2] + \mathbb{E}[\mathrm{Var}(V_a \mid M\mathbf{s}_t)].$$

Equality of the two risks forces the nonnegative conditional variance term to vanish almost surely. Consequently,

$$V_a = \mathbb{E}[V_a \mid M\mathbf{s}_t] \quad \text{a.s.}$$

The pair $(V_a, M\mathbf{s}_t)$ is jointly centered Gaussian because both variables are linear transformations of the Gaussian state. Its conditional mean is linear even when $M\Sigma_S M^\top$ is singular:

$$\mathbb{E}[V_a \mid M\mathbf{s}_t] = \beta_a^\top \Sigma_S M^\top (M\Sigma_S M^\top)^\dagger M\mathbf{s}_t,$$

where † denotes the Moore–Penrose inverse. Define

$$\ell_a^\top = \beta_a^\top \Sigma_S M^\top (M\Sigma_S M^\top)^\dagger.$$

Then $\beta_a^\top \mathbf{s}_t = \ell_a^\top M\mathbf{s}_t$ almost surely, so

$$\begin{aligned} 0 &= \mathbb{E}\Big[\big((\beta_a - M^\top \ell_a)^\top \mathbf{s}_t\big)^2\Big] \\ &= (\beta_a - M^\top \ell_a)^\top \Sigma_S (\beta_a - M^\top \ell_a). \end{aligned}$$

Positive definiteness of $\Sigma_S$ implies

$$\beta_a = M^\top \ell_a, \qquad \beta_a^\top \mathbf{s}_t = \ell_a^\top \mathbf{z}_t^c \quad \text{a.s.}$$

Since the argument holds for every $a \in \mathcal{A}$, it also yields

$$\ker(M) \subseteq \bigcap_{a \in \mathcal{A}} \ker(\beta_a^\top).$$

**Step 3: conclude recovery of the predictive subspace.** Steps 1 and 2 show that $\mathbf{z}_t^c$ removes the nuisance component $\mathbf{u}_t$ while preserving all linear directions of $\mathbf{s}_t$ that determine the potential-outcome means. Therefore, for every treatment level $a$,

$$\mathbb{E}[\mathrm{y}_{t+1}^a \mid \mathbf{s}_t] = \ell_a^\top \mathbf{z}_t^c \quad \text{a.s.}$$

This is the treatment-outcome predictive component needed by the outcome heads.

If the predictive subspace spans the full latent state, $\mathcal{S}_y = \mathbb{R}^{d_c}$, then

$$\bigcap_{a \in \mathcal{A}} \ker(\beta_a^\top) = \{0\}.$$

The inclusion

$$\ker(M) \subseteq \bigcap_{a \in \mathcal{A}} \ker(\beta_a^\top)$$

then gives $\ker(M) = \{0\}$. Since $M \in \mathbb{R}^{d_c \times d_c}$, $M$ is invertible. Setting $R^\top = M = \mathbf{W}^\top \mathbf{A}$ gives

$$\mathbf{z}_t^c = R^\top \mathbf{s}_t \quad \text{a.s.}$$

with $R$ invertible. □

**Scope of the recovery result.** The Gaussian-state assumption is what converts exact predictive sufficiency into preservation by linear functionals; positive-definite covariance alone would not justify this step. The rich-environment assumption requires at least $d_e + 1$ environments when $d_e > 0$, with nuisance mean differences spanning $\mathbb{R}^{d_e}$; the number of labels alone is insufficient. Full-state recovery additionally requires $\dim \mathcal{S}_y = d_c$. For scalar outcomes, $\dim \mathcal{S}_y \leq |\mathcal{A}|$, so binary treatments cannot support this full-recovery conclusion for a state dimension greater than two. These restrictions are not asserted for the learned 64-dimensional state or the benchmark environments. The theorem is a conditional population result, not an optimization guarantee or a proof of nonlinear causal disentanglement.

**Interpretation for EHR data.** Assumption 4 provides a tractable setting that illustrates the two learning principles used by DECAF: preserving treatment-outcome-relevant information while reducing reliance on environment-predictive variation. This structure matches a common source of shift in EHR data, where care units, cohorts, and hospitals may differ in treatment protocols, measurement frequency, documentation patterns, and patient-management practices. In such settings, separating stable treatment-response information from environment-related variation can improve robustness under policy or cohort shift. The linear-Gaussian analysis isolates this mechanism in a setting where recovery can be stated explicitly. In the empirical nonlinear setting, DECAF implements the same principle through outcome prediction, environment separation, and sufficiency regularization, and the experiments evaluate whether the resulting representation supports robust treatment-effect estimation and long-horizon forecasting.