Joint Point Process Model for Counterfactual Treatment–Outcome Trajectories Under Policy Interventions

Anonymous Author(s)
Affiliation
Address
email

Abstract

Policy makers need to predict the progression of an outcome before adopting a new treatment policy, which defines when and how a sequence of treatments affecting the outcome occurs in continuous time. Commonly, algorithms that predict interventional future outcome trajectories take a fixed sequence of future treatments as input. This excludes scenarios where the policy is unknown or a counterfactual analysis is needed. To handle these limitations, we develop a joint model for treatments and outcomes, which allows for the estimation of treatment policies and effects from sequential treatment–outcome data. It can answer interventional and counterfactual queries about interventions on treatment policies, as we show with a realistic semi-synthetic simulation study. This abstract is based on work that is currently under review (Anonymous).

1 Introduction

What policy should we adopt? In healthcare, for example, we observe patients' physiological markers (*outcomes*) changing over time. We want to affect these outcomes by actions (*treatments*) such as doses of a medicine. Sequences of outcomes and treatments are recorded as a time series. The choice of when to take what action constitutes the *policy*. To improve our policies, we must be able to assess their consequences: What is the effect of a given policy? What will be the effect of a change to a different policy? What would have happened if a patient had followed a different treatment policy? These questions correspond to observational, interventional, and counterfactual queries.

In high-risk domains such as public health and healthcare [Schulam and Saria, 2017, Bica et al., 2020], it is important to quantify risks and expectations accompanying the policy decision, as well as to evaluate the performance of past decisions [Oberst and Sontag, 2019, Tsirtsis and Gomez Rodriguez, 2020, Tsirtsis et al., 2021]. This requires estimating the causal effect of an intervention affecting the treatment policy on the outcome progression using a *causal* model.

Observed treatment–outcome data are always created by some existing policy; however, the policy is generally not recorded and may be known only implicitly through the observed data. Consequently, this link from past outcomes to future treatments is largely neglected in the sequential treatment–outcome literature, and the causal analysis is generally limited to a fixed sequence of treatment interventions set by hand or generated by a simplistic parametric model [Schulam and Saria, 2017, Lim et al., 2018, Bica et al., 2020, Seedat et al., 2022]. Such models cannot generalize beyond simulations to the analysis of realistic, alternative treatment policies in real-world applications. Also, evaluating treatment policies using counterfactual reasoning is not considered by most of the literature, which focuses on future progression.

Submitted to the 2022 NeurIPS Workshop on Gaussian Processes, Spatiotemporal Modeling, and Decision-making Systems. Do not distribute.

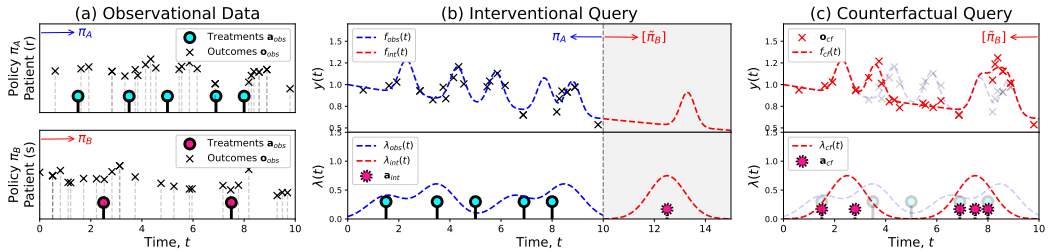

Figure 1: (a) Treatment–outcome data for two patients, following distinct policies π_A and π_B in the observation period $[0, 10]$. (b) The interventional query corresponds to how the outcome trajectory (r) will progress under a different policy π_B after the observation period $[0, 10]$ (shaded area). (c) The counterfactual query corresponds to how the outcome trajectory (r) would have progressed if the policy had been set to $\pi^{(r)} = \pi_B$ instead, in the observation period $[0, 10]$. Notice how the algorithm chooses to keep some of the observed treatments as counterfactual treatments $\mathbf{a_{cf}}$, where the counterfactual intensity is higher than the original observational intensity.

To address these limitations, we propose a joint treatment–outcome model. Our model can be learned from observational sequential treatment–outcome data (Figure 1(a)) and can estimate future and counterfactual progression. We show that an intervention on a treatment policy is equivalent to a stochastic intervention on sequential treatments, which we can model with our joint model, and use it to answer interventional (Figure 1(b)) and counterfactual (Figure 1(c)) queries.

2 Problem Definition

Consider an observational data set $\mathcal{D}$,

$$\mathcal{D} = \Big\{ \underbrace{\pi_{[0,T]}}_{\text{policy label}}, \underbrace{\{(t_i, m_i)\}_{i=1}^{N_a}}_{\text{treatments } \mathbf{a}}, \underbrace{\{(t_j, y_j)\}_{j=1}^{N_o}}_{\text{outcomes } \mathbf{o}} \Big\},$$

observed in the period $\mathcal{T} = [0, T]$. For notational simplicity, the data set is defined for a single individual. Our model can be trivially generalized to multiple individuals.

A policy label $\pi_{[0,T]} \in \Pi$ specifies a treatment intensity function $\lambda^*_\pi(t, m)$ in the interval $[0, T]$ that defines when and how a sequence of continuous-time treatments occur. We assume the data set $\mathcal{D}$ contains the policy label $\pi_{[0,T]}$, but its corresponding intensity function $\lambda^*_\pi(t, m)$ is unobserved. A treatment tuple $a_i = (t_i, m_i)$ consists of an arrival time t_i and a treatment mark m_i. An outcome tuple $o_j = (t_j, y_j)$ consists of a measurement time t_j and an outcome value y_j. The history $\mathcal{H}_{\leq t} = \{\pi_{\leq t}, \mathbf{a}_{\leq t}, \mathbf{o}_{\leq t}\}$ contains the information about the past policy $\pi_{\leq t}$, past treatments $\mathbf{a}_{\leq t} = \{(t_i, m_i) : t_i \leq t\}$ and past outcomes $\mathbf{o}_{\leq t} = \{(t_j, y_j) : t_j \leq t\}$.

We observe a continuous-time process $\mathbf{Y}_{\leq T} = \{y(\tau) : \tau \leq \mathcal{T}\}$ as outcome tuples $\mathbf{o}$ measured at times $\mathbf{t_o} = \{t_j\}_{j=1}^{N_o}$. To answer causal queries, we model the potential outcome trajectory $\mathbf{Y}_{>\tilde{\tau}}[\tilde{\pi}_{>\tilde{\tau}}]$, under an intervened policy specified by $\tilde{\pi}_{>\tilde{\tau}}$. When the intervention time $\tilde{\tau}$ is set to the end of the observation period $\tilde{\tau} = T$, we call the estimation task a *policy intervention*, as its computation requires access to the interventional distribution (Figure 1(b)):

$$P(\mathbf{Y}_{>T}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}). \tag{1}$$

Also, we can set the intervention time $\tilde{\tau}$ to the start of the observation period $\tilde{\tau} = 0$ and consider a hypothetical scenario under an alternative treatment policy specified by $\tilde{\pi}_{\leq T}$. We call this estimation task a *policy counterfactual*, as its computation requires access to the counterfactual distribution (Figure 1(c)):

$$P(\mathbf{Y}_{\leq T}[\tilde{\pi}_{\leq T}] \mid \mathcal{H}_{\leq T}). \tag{2}$$

Under a set of causal assumptions defined in Appendix C, we show in Appendix D that (i) the potential outcome trajectory $\mathbf{Y}_{>T}[\tilde{\pi}_{>T}]$ is equivalent to the potential outcome trajectory under a sequence of

stochastic interventions on treatments and (ii) it is identified using two statistical quantities [Pearl, 2009, Hernán and Robins, 2010]. To estimate these statistical quantities from the observational data, we propose a joint treatment–outcome model.

3 Treatment–Outcome Model

Our joint model is a combination of two conditional intensity functions: (i) treatment intensity: $\lambda_\pi^*(t, m) = \lambda_\pi^*(t)p^*(m \mid t)$ and (ii) outcome intensity: $\lambda_o^*(t, y) = \lambda_o^*(t)p^*(y \mid t)$. We assume the measurement times $\mathbf{t_o}$ of the outcomes are given, which is valid for example when the data are collected through automated patient monitoring in healthcare. Then, the joint distribution for the data set $\mathcal{D}$ can be written in terms of two intensity functions as follows [Daley and Vere-Jones, 2003, Rasmussen, 2011]:

$$p(\mathcal{D}) = \prod_{i=1}^{I} \underbrace{\lambda_\pi^*(t_i)p^*(m_i \mid t_i)}_{\text{Treat. Intensity}} \prod_{t_j \in \mathbf{t_o}} \underbrace{p^*(y_j \mid t_j)}_{\text{Out. Model}} \times \exp(-\Lambda), \tag{3}$$

with the integral term $\Lambda = \int_{\mathcal{T}} \lambda_\pi^*(\tau)d\tau$.

3.1 Treatment Intensity

We model the treatment time intensity $\lambda_\pi^*(\tau)$ using a constant baseline β_0 and three functions with GP priors, $g_b, g_a^*, g_o^* \sim \mathcal{GP}$. The latent-state function g_b models the baseline intensity. The regressive components g_a^* and g_o^* model the dependence on past treatments and outcomes respectively [Liu and Hauskrecht, 2019]. The treatment intensity $\lambda_\pi^*(\tau)$ is defined as follows:

$$\lambda_\pi^*(\tau) = \Big(\underbrace{\beta_0}_{\substack{\text{PP}\\ \text{Baseline}}} + \underbrace{g_b(\tau)}_{\substack{\text{NHPP}\\ \text{Baseline}}} + \underbrace{g_a^*(\tau)}_{\substack{\text{Treat.}\\ \text{Effect}}} + \underbrace{g_o^*(\tau)}_{\substack{\text{Out.}\\ \text{Effect}}} \Big)^2.$$

The model and kernel definitions are detailed in Appendix F.1-2.

3.2 Outcome Model

We model the outcome trajectory $\mathbf{Y} = \{y(\tau) : \tau \in \mathbb{R}_{\geq 0}\}$ over time τ, combining three independent components: (i) a baseline progression, (ii) treatment effects and (iii) a noise variable [Schulam and Saria, 2017, Xu et al., 2016, Zhang et al., 2020]:

$$y(\tau) = \underbrace{f_b(\tau)}_{\text{Baseline}} + \underbrace{f_a(\tau; \mathbf{a})}_{\text{Treatment Effect}} + \underbrace{\epsilon(\tau)}_{\text{Noise}}. \tag{4}$$

The baseline progression and the treatment effect functions are modeled by GP priors, with an independent Gaussian noise $\epsilon(\tau) \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The model and kernel definitions are detailed in Appendix G.

4 Experiments

We evaluate our model on two causal inference tasks: (i) the policy intervention (Equation 1) and (ii) the policy counterfactual (Equation 2), by setting up a realistic semi-synthetic simulation scenario.

We fit our joint model to a challenging real-world data set on meal–blood glucose dynamics [Zhang et al., 2020, Wyatt et al., 2021] to obtain the ground-truth data generators. The ground-truth models are used to simulate samples from observational, interventional and counterfactual distributions. Simulated patients are divided into two policy groups $\{\pi_A, \pi_B\}$, representing different treatment policies of different hospitals, countries, etc. The details of the simulation study are presented in Appendix J.

We define three joint estimation models: OBS-EST, INT-EST and CF-EST. OBS-EST is trained on the observational data of each individual to generate predictions. INT-EST adjusts predictions of OBS-EST by accounting for the fact that the treatments are generated by the estimated policy for another individual, as a consequence of a policy intervention. CF-EST additionally conditions

Table 1: DACC results for two policy interventions $\{[\tilde{\pi}_{>T} = \pi_A], [\tilde{\pi}_{>T} = \pi_B]\}$ over 10 runs. The observed policy is $\pi_{[0,T]} = \pi_A$. DACC closer to 50% is better, as it suggests estimated trajectories are inseparable from ground-truth trajectories.

	$[\tilde{\pi} = \pi_A]$	$[\tilde{\pi} = \pi_B]$
JOINT MOD.	DACC ↓	DACC ↓
OBS-EST	$51.8 \pm 2.8\%$	$86.4 \pm 2.5\%$
INT-EST	$51.8 \pm 2.8\%$	$51.8 \pm 2.8\%$
INT-ORACLE	$50.3 \pm 2.3\%$	$50.3 \pm 2.3\%$

Table 2: DACC results for the policy counterfactual in the observed period $[0, T]$ with the policy intervention $[\tilde{\pi}_{[0,T]} = \pi_B]$ over 10 runs. The observed policy is $\pi_{[0,T]} = \pi_A$. DACC closer to 50% is better, as it suggests estimated trajectories are inseparable from ground-truth trajectories.

	$[\tilde{\pi} = \pi_B]$
JOINT MOD.	DACC ↓
INT-EST	$90.1 \pm 4.1\%$
CF-EST	$60.8 \pm 2.2\%$
CF-ORACLE	$51.8 \pm 2.7\%$

predictions of INT-EST with the posterior of the individual's noise terms. We denote ground-truth versions of these models as OBS-ORACLE, INT-ORACLE and CF-ORACLE, which represent the performance of the estimated models if infinite training data were available.

To measure how similar predicted trajectories are to samples from the ground-truth distribution, we train discriminators. Ideally, for samples of the same distribution, predicted trajectories should be inseparable from ground-truth trajectories, leading to a 50% discriminator accuracy (DACC).

For the policy intervention task (Table 1), we see that the INT-EST model is able to sample observational and interventional trajectories close to ground-truth distributions when the intervention policy is (i) same as the observed policy $[\tilde{\pi} = \pi_A]$ and (ii) different from the observed policy $[\tilde{\pi} = \pi_B]$, while the OBS-EST model fails in the latter case. For the policy counterfactual task (Table 2), we see the INT-EST model fails to sample counterfactual trajectories close to the ground-truth counterfactual distribution $[\tilde{\pi} = \pi_B]$, as it does not take the individual's noise posterior into account. On the other hand, the CF-EST model is able to sample counterfactual trajectories close to the ground-truth counterfactual distribution.

5 Conclusion

To study what happens if the (possibly implicit) treatment policy of one individual (hospital, country, ...) is or had been adopted by another individual, we proposed a model that jointly considers sequences of treatments and outcomes of each individual. Theoretically, we showed that an intervention on a treatment policy is equivalent to a sequence of stochastic interventions on treatments, whose potential outcomes can be estimated from observational data with the joint model. In a semi-synthetic experiment, we demonstrated that the joint model can answer causal queries about the interventional and counterfactual distributions of the outcome after an intervention on the treatment policy.

References

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*, 2020.

Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer, 2003.

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

Bryan Lim, Ahmed M Alaa, and Mihaela van der Schaar. Forecasting treatment responses over time using recurrent marginal structural networks. *advances in neural information processing systems*, 31, 2018.

Siqi Liu and Milos Hauskrecht. Nonparametric regressive point processes based on conditional gaussian processes. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 1064–1074, 2019.

Michael Oberst and David Sontag. Counterfactual off-policy evaluation with gumbel-max structural causal models. In *International Conference on Machine Learning*, pages 4881–4890. PMLR, 2019.

Judea Pearl. *Causality*. Cambridge university press, 2009.

Jakob Gulddahl Rasmussen. Temporal point processes: the conditional intensity function. *Lecture Notes, Jan*, 2011.

Peter Schulam and Suchi Saria. Reliable decision support using counterfactual models. *Advances in neural information processing systems*, 30, 2017.

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.

Stratis Tsirtsis and Manuel Gomez Rodriguez. Decisions, counterfactual explanations and strategic behavior. *Advances in Neural Information Processing Systems*, 33:16749–16760, 2020.

Stratis Tsirtsis, Abir De, and Manuel Rodriguez. Counterfactual explanations in sequential decision making under uncertainty. *Advances in Neural Information Processing Systems*, 34:30127–30139, 2021.

Patrick Wyatt, Sarah E Berry, Graham Finlayson, Ruairi O'Driscoll, George Hadjigeorgiou, David A Drew, Haya Al Khatib, Long H Nguyen, Inbar Linenberg, Andrew T Chan, et al. Postprandial glycaemic dips predict appetite and energy intake in healthy individuals. *Nature metabolism*, 3(4): 523–529, 2021.

Yanbo Xu, Yanxun Xu, and Suchi Saria. A bayesian nonparametric approach for estimating individualized treatment-response curves. In *Machine learning for healthcare conference*, pages 282–300. PMLR, 2016.

Guangyi Zhang, Reza A Ashrafi, Anne Juuti, Kirsi Pietiläinen, and Pekka Marttinen. Errors-in-variables modeling of personalized treatment-response trajectories. *IEEE Journal of Biomedical and Health Informatics*, 25(1):201–208, 2020.

Supplementary Material: Joint Point Process Model for Counterfactual Treatment–Outcome Trajectories Under Policy Interventions

Anonymous Author(s)
Affiliation
Address
email

A Further Related Work

Treatment Responses for Sequential Treatment–Outcome Data. In a static treatment–outcome setup, the goal is to estimate the causal effect of a single treatment on a single outcome [Pearl, 2009]. The methods for the static setup does not extend to the sequential treatment–outcome setup directly, due to the time-varying confounding [Hernán and Robins, 2010]. For the sequential treatment–outcome setup, a number of methods such as g-computation [Robins, 1986, 1987], structural nested models [Robins, 1992] and marginal structural models [Robins et al., 2000] have been proposed.

Our work is in line with the g-computation formula [Robins, 1986], as we also model the joint distribution directly by factorizing it in the temporal order. In this direction, a large number of methods model the outcome trajectory as the sum of (i) a counterfactual baseline that captures the no-treatment case and (ii) an additive treatment response function [Schulam and Saria, 2015, Xu et al., 2016, Soleimani et al., 2017, Zhang et al., 2020, Cheng et al., 2020]. Most of these work propose a Bayesian non-parametric baseline and a parametric response function [Schulam and Saria, 2015, Xu et al., 2016, Soleimani et al., 2017, Zhang et al., 2020]. In our work, we combine the non-parametric baseline with a non-parametric response function as in Cheng et al. [2020]. However, these methods do not model treatments and they assume a fixed sequence of treatment interventions are given.

Treatment Models in Continuous Time. Continuous-time treatments can be viewed as discrete events in a continuous space, which can be modeled appropriately by a temporal point process (TPP). In this direction, temporal point processes have been investigated from a causal inference perspective by a large number of work [Lok, 2008, Didelez et al., 2012, Schulam and Saria, 2017, Gao et al., 2021, Aalen et al., 2020, Ryalen et al., 2020, Hua et al., 2021].

Lok [2008] estimates the causal effect of interventions on continuous-time treatments, by extending structural nested models using a martingale approach. The methods proposed by Didelez et al. [2012], Aalen et al. [2020] and Ryalen et al. [2020] consider interventions affecting continuous-time events on a survival analysis scenario, where the TPP terminates after a single survival/death event. Gao et al. [2021] investigates the average treatment effect between multivariate temporal event pairs in continuous time. Similar to our work, Schulam and Saria [2017] estimates interventional outcome (mark) trajectories affected by continuous-time treatments, however, it assumes a fixed sequence of treatment interventions. Instead, we model continuous-time treatment policies to consider more realistic, alternative scenarios.

From a modeling perspective, Hua et al. [2021] is the closest to our work, where they propose a Bayesian joint model for treatment–outcome tuples, however, they aim to find an optimal treatment policy for continuous-time treatments, while our goal is to estimate interventional and counterfactual outcome trajectories resulting from interventions on treatment policies. Besides, they use simple,

Submitted to the 2022 NeurIPS Workshop on Gaussian Processes, Spatiotemporal Modeling, and Decision-making Systems. Do not distribute.

parametric models as treatment and outcome models, tailored for their use case, which does not generalize to other problem setups.

Counterfactual Reasoning. Counterfactual reasoning has raised interest recently in the machine learning literature for evaluating and explaining model predictions [Oberst and Sontag, 2019, Tsirtsis and Gomez Rodriguez, 2020, Tsirtsis et al., 2021, Abid et al., 2022]. These works consider discrete-time action–outcome setups. On the other hand, the works by [Aalen et al., 2020, Ryalen et al., 2020] and [Noorbakhsh and Rodriguez, 2021] consider counterfactuals of continuous-time events using temporal point processes. However, [Aalen et al., 2020] and [Ryalen et al., 2020] target a survival analysis setup as discussed above. In our work, we extend the counterfactual sampling algorithm of independent Poisson processes, provided in [Noorbakhsh and Rodriguez, 2021], to history-dependent point processes. In addition, our main target is the outcome (mark) trajectories that are affected by treatment policies through continuous-time treatments, while [Noorbakhsh and Rodriguez, 2021] investigates the causal effect on continuous-time event occurrences as target outcomes.

B Intervention Types

In practice, an intervention on a variable X remove the structural equation $X := f_x(pa(X), U_X)$. Intervention types differ with respect to the way they set the intervened value, $\tilde{x}$ value.

Atomic intervention. Sets X to a static value $\tilde{x}$. For example, under identifiability conditions:

$$p(Y \mid do(X = \tilde{x})) = \sum_z p(Y \mid X = \tilde{x}, Z = z)p(Z = z).$$

Conditional intervention. Sets X by a deterministic functional relationship $\tilde{h}(pa(X))$. For example, under identifiability conditions:

$$\begin{aligned} p(Y \mid do(X = \tilde{h}(z))) &= \sum_z p(Y \mid do(X = \tilde{h}(z)), Z = z)p(Z = z \mid do(X = \tilde{h}(z))) \\ &= \sum_z p(Y \mid x, z)|_{x=\tilde{h}(z)} p(z). \end{aligned}$$

Stochastic (conditional) intervention. Sets X by a stochastic relationship $\tilde{p}(X \mid Z)$. Given Z, the intervention $do(X = \tilde{x})$ occurs with probability $\tilde{p}(X \mid Z)$. Therefore, the effect of an atomic intervention $do(X = \tilde{x})$ is averaged over all possible $\tilde{x} \sim \tilde{p}(x|z)$:

$$\begin{aligned} &p(Y \mid do(X \sim \tilde{p}(X \mid Z))) \\ &\quad = \sum_{\tilde{x}} \sum_z p(Y \mid do(X = \tilde{x}), z = z)\tilde{p}(X = \tilde{x} \mid Z = z)p(Z = z) \\ &\quad = \sum_{\tilde{x}} \sum_z p(Y \mid x, z)\tilde{p}(\tilde{x} \mid z)p(z). \end{aligned}$$

C Causal Assumptions

Consistency, no-unmeasured confounding (NUC) and positivity are standard causal assumptions required for the identifiability of a static causal effect [Pearl, 2009, Hernán and Robins, 2010, Schulam and Saria, 2017, Bica et al., 2020].

Assumption 1: Consistency. Let A be an action variable and Y be an outcome variable. The potential outcome $Y[a]$ under an action $a \in \mathcal{A}$ is consistent with its factual outcome, provided that the outcome Y is observed under the action $A = a$: $P(Y[a] \mid A = a) = P(Y \mid A = a)$.

Assumption S1: No-Unmeasured Confounding (NUC). Let A be an action variable, Y be an outcome variable and X be a set of observed variables. The potential outcome $Y[a]$ is independent of the action A conditioned on X, if the action assignment is done at random given X: $Y[a]A \mid X$.

Assumption S2: Positivity (Overlap). Let A be an action variable and X be a set of variables rendering the causal effect of A identifiable. For a well-defined causal effect, the action must be probable conditioned on X: $P(A = a \mid X) > 0$.

For the continuous-time sequential treatment–outcome setup, Schulam and Saria [2017] propose the continuous-time NUC assumption. Assumptions $\{3, 4\}$ are repeated below for completeness.

Assumption 2: Continuous-Time Positivity. The conditional treatment intensity $\lambda^*_\pi(t, m)$ is non-zero for all intervention times t and all intervention marks m, given any intervention history $\mathcal{H}_{<t}$: $\lambda^*_\pi(t, m) > 0, \forall t \in \mathbb{R}, \forall m \in \mathcal{K}_a$.

Assumption 3: Continuous-time NUC. Let sequential treatments $\mathbf{a}$ occur at a discrete set of time points on a given interval, characterized by the conditional intensity function $\lambda^*_\pi(t, m)$. The conditional treatment intensity $\lambda^*_\pi(t, m)$ of a treatment $a = (t, m)$, is independent of the potential outcome trajectory $\mathbf{Y}_{>t}[\tilde{\mathbf{a}}]$, conditioned on the past history $\mathcal{H}_{<t}$, $\forall t \in \mathbb{R}_{\geq 0}$[1].

Assumption 4: Fully-mediated policy effect. Conditioned on the past history $\mathcal{H}_{<\tau}$, the causal effect of the policy $\pi_{>\tau}$ on the outcome trajectory $\mathbf{Y}_{>\tau}$ is fully mediated through sequential treatments $\mathbf{a}_{>\tau}$:

$$P(Y_{>\tau}[\tilde{\pi}_{>\tau}, \tilde{\mathbf{a}}_{>\tau}] \mid \mathcal{H}_{\leq\tau}) = P(Y_{>\tau}[\tilde{\mathbf{a}}_{>\tau}] \mid \mathcal{H}_{\leq\tau}).$$

D Causal Identifiability

We evaluate the potential outcome trajectory $\mathbf{Y}_{>T}[\tilde{\pi}_{>T}]$ at a discrete set of ordered query points $\mathbf{q} = \{q_1, ..., q_m : q_i > T, \forall i \in 1, ..., m\}$: $\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}] = \{y(q_k)[\tilde{\pi}_{>T}] : q_k \in \mathbf{q}\}$, similar to Schulam and Saria [2017]. We denote the outcome value $y(q_k)$ at time q_k by $Y_{q_k} = y(q_k)$. Under the NUC Assumption (Assumption S1), the assignment of the target policy label $\tilde{\pi}$ is independent of the potential outcome query $\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}]$:

$$P(\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}) = P(\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}] \mid \tilde{\pi}_{>T}, \mathcal{H}_{\leq T})$$

As the causal graph $\mathcal{G}$ suggests, sequential actions $\mathbf{a}$ act as mediators in the time direction for the causal effect of the policy specified by $\tilde{\pi}$ on the outcome query $\mathbf{Y}_{\mathbf{q}}$. To take the action mediation into account, we average over possible treatment effects:

$$P(\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}] \mid \tilde{\pi}_{>T}, \mathcal{H}_{\leq T}) = \sum_{\mathbf{a}_{>T}} P(\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}], \mathbf{a}_{>T} \mid \tilde{\pi}_{>T}, \mathcal{H}_{\leq T}).$$

An intermediate outcome Y_{q_k} can act as a mediator for the causal effect of past actions $\mathbf{a}_{<q_k}$ on the future outcome $Y_{q_{k+1}}$, while acting as a confounder for some future actions $\mathbf{a}_{[q_k, q_{k+1})}$ and the future outcome $Y_{q_{k+1}}$, causing a phenomenon known as time-varying confounding [Robins, 1986, Pearl and Robins, 1995]. To eliminate the time-varying confounding effect, we first factorize the counterfactual query $\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}]$ and sequential actions $\mathbf{a}_{>T}$ in time-order:

$$P(\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}) = \sum_{\mathbf{a}} \prod_{k=1}^{m} \Big(P(Y_{q_k}[\tilde{\pi}_{>T}] \mid \mathbf{a}_{<q_k}, \mathbf{Y}_{\leq q_{k-1}}[\tilde{\pi}_{>T}], \tilde{\pi}_{>T}, \mathcal{H}_{\leq T}) \\ P(\mathbf{a}_{[q_{k-1}, q_k)} \mid \mathbf{a}_{<q_{k-1}}, \mathbf{Y}_{\leq q_{k-1}}[\tilde{\pi}_{>T}], \tilde{\pi}_{>T}, \mathcal{H}_{\leq T}) \Big), \quad (1)$$

similar to the g-estimation formula [Robins, 1986]. In both terms, we condition on the policy label $\tilde{\pi}_{>T}$. By consistency assumption (Assumption 1), the potential outcome $\mathbf{Y}_{\leq q_{k-1}}[\tilde{\pi}_{>T}]$ is equal to the factual outcome when we observe the policy specified by $\tilde{\pi}_{>T}$:

$$P(\mathbf{Y}_{\mathbf{q}}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}) = \sum_{\mathbf{a}} \prod_{k=1}^{m} \Big(P(Y_{q_k}[\tilde{\pi}_{>T}] \mid \mathbf{a}_{<q_k}, \mathbf{Y}_{\leq q_{k-1}}, \tilde{\pi}_{>T}, \mathcal{H}_{\leq T}) \\ P(\mathbf{a}_{[q_{k-1}, q_k)} \mid \mathbf{a}_{<q_{k-1}}, \mathbf{Y}_{\leq q_{k-1}}, \tilde{\pi}_{>T}, \mathcal{H}_{\leq T}) \Big). \quad (2)$$

[1] In fact, the history $\mathcal{H}_{<t}$ can be extended to include an outcome variable that occurs at time t, if it exists: $\mathcal{H}_{\leq t} \cup \{o_t\}$, assuming instantaneous effects from outcomes to treatments: $o \to a$. However, it would clutter the notation and in practice, the probability of an outcome $o = (t, y)$ to occur at time t is 0. Therefore, assuming no instantaneous effects between variables can be made without any practical effects.

For both the outcome variable Y_{q_k} and actions $\mathbf{a}_{[q_{k-1},q_k)}$, the set of variables $\{\mathbf{a}_{<q_{k-1}}, \mathbf{Y}_{\leq q_{k-1}}, \mathcal{H}_T\}$ is a valid history until the query point q_{k-1}: $\mathcal{H}_{q_{k-1}} = \mathcal{H}_{\leq q_{k-1}} \setminus \{a_{q_{k-1}}\} = \{\mathbf{a}_{<q_{k-1}}, \mathbf{Y}_{\leq q_{k-1}}, \mathcal{H}_T\}$. Hence, the Equation 1 can be simplified as follows:

$$P(\mathbf{Y_q}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}) = \sum_{\mathbf{a}} \prod_{k=1}^{m} \Big(\underbrace{P(Y_{q_k}[\tilde{\pi}_{>T}] \mid \mathbf{a}_{[q_{k-1},q_k)}, \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}})}_{\text{Outcome Factor}} \underbrace{P(\mathbf{a}_{[q_{k-1},q_k)} \mid \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}})}_{\text{Treatment Factor}} \Big), \quad (3)$$

Under Assumptions $\{1, S1\}$, the treatment factor term is equivalent to the potential outcome $\mathbf{a}_{[q_{k-1},q_k)}[\tilde{\pi}_{>T}]$:

$$\begin{aligned} P(\mathbf{a}_{[q_{k-1},q_k)} \mid \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}}) &= P(\mathbf{a}_{[q_{k-1},q_k)}[\tilde{\pi}_{>T}] \mid \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}}) && \text{A.1} \\ &= P(\mathbf{a}_{[q_{k-1},q_k)}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{q_{k-1}}) && \text{A.S1} \end{aligned} \quad (4)$$

Conditioned on the past history $\mathcal{H}_{q_{k-1}}$, the causal effect of the policy specified by $\tilde{\pi}$ on the outcome value Y_{q_k} is fully mediated through sequential treatments $\mathbf{a}_{[\mathbf{q_{k-1}},\mathbf{q_k})}$. Using Assumptions $\{1, 2, 3, 4\}$, the outcome factor is equivalent to the potential outcome $Y_{q_k}[\tilde{\mathbf{a}}_{[q_{k-1},q_k)}]$:

$$\begin{aligned} P(Y_{q_k}[\tilde{\pi}_{>T}] \mid &\mathbf{a}_{[q_{k-1},q_k)}, \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}}) \\ &= P(Y_{q_k} \mid \mathbf{a}_{[q_{k-1},q_k)}, \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}}) && \text{A.1} \\ &= P(Y_{q_k} \mid \mathbf{a}_{[q_{k-1},q_k)}, \mathcal{H}_{q_{k-1}}) && \text{A.4} \\ &= P(Y_{q_k}[\mathbf{a}_{[q_{k-1},q_k)}] \mid \mathbf{a}_{[q_{k-1},q_k)}, \mathcal{H}_{q_{k-1}}) && \text{A.1} \\ &= P(Y_{q_k}[\mathbf{a}_{[q_{k-1},q_k)}] \mid \mathcal{H}_{q_{k-1}}) && \text{A.3} \end{aligned} \quad (5)$$

Plugging in the potential outcome representation of the treatment factor in Equation 4 the outcome factor in Equation 5 to the Equation 3, the potential query can be considered as a sequence of stochastic (conditional) interventions on treatment variables:

$$P(\mathbf{Y_q}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}) = \sum_{\tilde{\mathbf{a}}} \prod_{k=1}^{m} P(Y_{q_k}[\tilde{\mathbf{a}}_{[q_{k-1},q_k)}]] \mid \mathcal{H}_{q_{k-1}}) P(\tilde{\mathbf{a}}_{[q_{k-1},q_k)}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{q_{k-1}}). \quad (6)$$

Under Assumptions $\{1, 2, 3\}$, a sequence of stochastic (conditional) interventions on treatments is identified [Seedat et al., 2022, Didelez, 2015]:

$$P(\mathbf{Y_q}[\tilde{\pi}_{>T}] \mid \mathcal{H}_{\leq T}) = \sum_{\tilde{\mathbf{a}}} \prod_{k=1}^{m} P(Y_{q_k} \mid \tilde{\mathbf{a}}_{[q_{k-1},q_k)}, \mathcal{H}_{q_{k-1}}) P(\tilde{\mathbf{a}}_{[q_{k-1},q_k)} \mid \tilde{\pi}_{>T}, \mathcal{H}_{q_{k-1}}). \quad (7)$$

As discussed in Appendix I.2, the posterior distributions of the noise variables for the counterfactual temporal point processes are identifiable, as they satisfy counterfactual stability. Assuming the posterior distributions of the outcome noise variables as identifiable, the counterfactual query is also identifiable[2]

E Nonparametric Modeling approaches for TPPs

TPPs can be modeled by two nonparametric approaches Liu and Hauskrecht [2019]: (i) latent-state point processes Møller et al. [1998], Adams et al. [2009] and (ii) regressive point processes Hawkes [2018], Liu and Hauskrecht [2019].

[2]The counterfactual stability condition does not follow for the outcome noise variables, since they are continuous. Even though the posterior of the outcome noise variables do not satisfy the counterfactual stability, generally the assumption that the noise posteriors are identifiable is implicitly made for the inferences regarding the interventional distribution in the literature, e.g. the estimated interventional trajectories in Schulam and Saria [2017] use a GP latent function with Gaussian noise variables similarly.

E.1 Latent-state point processes

Latent-state point processes, e.g. a log/sigmoidal Gaussian Cox process [Møller et al., 1998, Adams et al., 2009] and a variational Bayes point process (VBPP) [Lloyd et al., 2015], model the temporal dependence of the intensity indirectly, in the form of smoothness assumptions encoded by a latent (GP) prior [Rasmussen, 2003]. For example, the VBPP model defines a latent GP prior $f \sim \mathcal{GP}$ with a squared-exponential (SE) kernel over the observation period $\mathcal{T} = [0, T]$, which is passed through a square transformation to obtain the conditional intensity function $\lambda(\tau)$. For a vector of N time points $\boldsymbol{\tau} \in \mathbb{R}^N_{\geq 0}$, the model definition takes the following form:

$$
\begin{aligned}
&f(\boldsymbol{\tau}) \sim \mathcal{GP}(\mu(\boldsymbol{\tau}), k(\boldsymbol{\tau}, \boldsymbol{\tau'})), && \boldsymbol{\tau}, \boldsymbol{\tau'} \in \mathbb{R}^N_{\geq 0} \\
&\lambda(\boldsymbol{\tau}) = f(\boldsymbol{\tau})^2.
\end{aligned}
$$

The latent function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ takes an absolute time $\tau \in \mathcal{T}$ as input and outputs the latent rate for the same time point. The choice of the SE kernel assumes a smooth latent function, which implicitly enforces some temporal dependence of nearby intensity values. For example, the intensity values for points $\{t_i\}_{i=1}^4$ in Figure 1 is given by:

$$
\lambda^* \left(\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} \right) = f \left(\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} \right)^2 .
$$

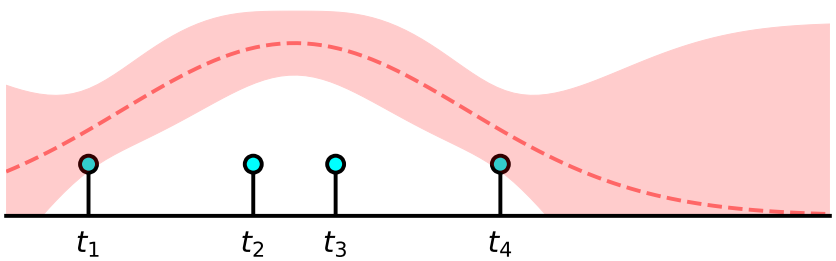

Figure 1: An example of the conditional intensity function of a latent-state point process model. The conditional intensity takes absolute time values as input and gives an intensity value as output.

E.2 Regressive Point Processes

In contrast, regressive point processes model the temporal dependence directly [Hawkes, 2018, Liu and Hauskrecht, 2019]. For example, the conditional GP regressive point process (CGPRPP) model defines a latent GP prior $f \sim \mathcal{GP}$, which is conditioned on past events. The latent function $f : \mathbb{R}^D_{\geq 0} \to \mathbb{R}$ takes relative times of the last D events relative to a future time point τ as input. For example, let a query time τ and let $(t_{i-(D-1)}, \ldots, t_{i-1}, t_i)$ be a list of arrival times of the last D events. To calculate the intensity value $\lambda^*(\tau)$, the input takes the following form:

$$
(\tau - t_{i-(D-1)}, \ldots, \tau - t_{i-1}, \tau - t_i).
$$

Let $\Delta\tau_d$ be the relative time between τ and d^{th} last event: $\Delta\tau_d = \tau - t_{i-d}$. Then, the input can be represented in a compact form: $\{\Delta\tau_d\}_{d=0}^{D-1}$. Further, let $\boldsymbol{\Delta\tau}$ denote the input set $\{\Delta\tau_d\}_{d=0}^{D-1}$ as a shorthand notation.

CGPRPP model regresses relative times $\{\Delta\tau_d\}_{d=1}^{D}$ on the latent intensity $f(\tau)$, which is passed through a square transformation to obtain the conditional intensity value $\lambda^*(\tau)$. For a vector of N time points $\boldsymbol{\tau} \in \mathbb{R}^N_{\geq 0}$, the model definition takes the following form:

$$
\begin{aligned}
&f(\boldsymbol{\tau}) \sim \mathcal{GP}(\mu(\boldsymbol{\Delta\tau}), k(\boldsymbol{\Delta\tau}, \boldsymbol{\Delta\tau'}), && \boldsymbol{\Delta\tau}, \boldsymbol{\Delta\tau'} \in \mathbb{R}^{N \times D}_{\geq 0} \\
&\lambda(\boldsymbol{\tau}) = f(\boldsymbol{\tau})^2.
\end{aligned}
$$

Unlike a latent-state formulation, the regressive latent function uses relative times as input instead of absolute times and models the dependence on past events explicitly. As an example, the intensity values for points $\{t_i\}_{i=1}^4$ in Figure 2, with $D = 2$, is given by:

$$\lambda^* \left(\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} \right) = f \left(\begin{bmatrix} \infty & \infty \\ t_2 - t_1 & \infty \\ t_3 - t_2 & t_3 - t_1 \\ t_4 - t_3 & t_4 - t_2 \end{bmatrix} \right)^2$$

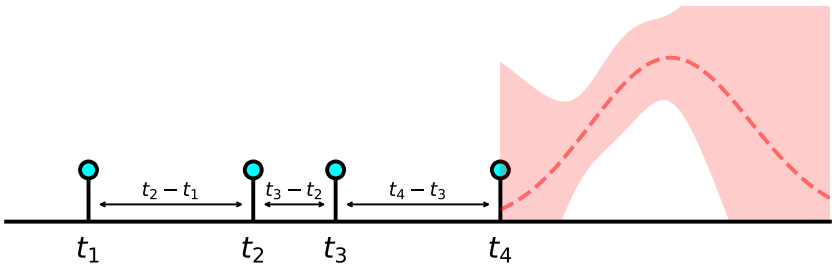

Figure 2: An example of the conditional intensity function of a regressive point process model, with $D = 2$. The conditional intensity takes relative time values to last $D = 2$ points as input and gives an intensity value as output.

F Treatment Model Details

F.1 Model Definition

We model the treatment time intensity $\lambda^*(\tau)$ as the square transformation of a latent function, which is a sum of a constant scalar β_0 and three time-dependent functions with GP priors, $g_b, g_a^*, g_o^* \sim \mathcal{GP}$. The latent-state function $g_b : \mathbb{R}_{\geq 0} \to \mathbb{R}$ models the baseline intensity. The regressive component $g_a^* : \{\mathbb{R}_{\geq 0} \times \mathcal{K}_a\}^{Q_a} \to \mathbb{R}$ models the dependence on past treatments by taking the last Q_a treatment tuples as input. Similarly, the regressive component $g_o^* : \{\mathbb{R}_{\geq 0} \times \mathcal{K}_o\}^{Q_o} \to \mathbb{R}$ models the dependence on past outcomes by taking the last Q_o outcome tuples as input.

The treatment time intensity $\lambda^*(\tau)$ is as follows:

$$\lambda^*(\tau) = \Big(\underbrace{\beta_0}_{\substack{\text{PP} \\ \text{baseline}}} + \underbrace{g_b(\tau)}_{\substack{\text{NHPP} \\ \text{Baseline}}} + \underbrace{g_a^*(\tau)}_{\substack{\text{Treatment} \\ \text{Effect}}} + \underbrace{g_o^*(\tau)}_{\substack{\text{Outcome} \\ \text{Effect}}} \Big)^2.$$

F.2 Kernel Definition

Regressive components g_a^* and g_o^* explicitly model how the intensity depends on the past history. They take relative times to the last Q_a treatments and Q_o outcomes as input, while the baseline function g_b takes the absolute time. For example, let τ be a query time. Let $(t_{i-(Q_a-1)}, \ldots, t_{i-1}, t_i)$ be the list of arrival times of the last Q_a treatments, before time τ. Similarly, let $(t_{j-(Q_o-1)}, \ldots, t_{j-1}, t_j)$ be the list of arrival times of the last Q_o outcomes, before time τ. Then, three function components has the following $1 + 2 * Q_a + 2 * Q_o$ dimensional vector as input:

$$\tau, \underbrace{\ldots, (\tau - t_i, m_i)}_{\text{last } Q_a \text{ Treatments}}, \underbrace{\ldots, (\tau - t_j, y_j)}_{\text{last } Q_o \text{ Outcomes}}$$

To represent the information regarding the last $Q_a + Q_o$ events, we define retrieval functions. Let $r_{k_i,t} : \mathbb{R} \to \mathbb{R}$ denote a retrieval function that takes time τ as input and outputs the time of the ith last tuple of type $k \in \{a, o\}$. For example, the retrieval function $r_{a_1,t}(\tau)$ outputs the time of the last action occurred before time τ. Let $\Delta_{k_i,t} : \mathbb{R} \to \mathbb{R}$ denote a retrieval function that returns the relative time to the ith last tuple of type $k \in \{a, o\}$: $\Delta_{k_i,t}(\tau) = \tau - q_{k_i,t}$. Similarly, $r_{k_i,m} : \mathbb{R} \to \mathcal{K}_{k_i}$

denotes a retrieval function that takes time τ as input and provides the mark $m \in \mathcal{K}_k$ of ith last tuple of type $k \in \{a, o\}$. Let $r_{k_i} : \mathbb{R} \to \mathcal{K}_k \times \mathbb{R}$ denote a retrieval function that outputs the mark and the relative time of the ith last tuple of type $k \in \{a, o\}$: $r_{k_i}(\tau) = (r_{k_i,m}(\tau), \Delta_{k_i,t}(\tau))$. For a query time τ, the retrieval function $r(\tau) : \mathbb{R}_{\geq 0} \to \mathcal{X}$ returns the overall input vector:

$$r(\tau) = \{\tau, \underbrace{r_{a_1}(\tau), \ldots, r_{a_{Q_a}}(\tau)}_{\text{last } Q_a \text{ Treatments}}, \underbrace{r_{o_1}(\tau), \ldots, r_{o_{Q_o}}(\tau)}_{\text{last } Q_o \text{ Outcomes}}\},$$

where $\mathcal{X}$ denotes the input domain $\mathcal{X} = \mathbb{R}_{\geq 0} \times \{\mathbb{R}_{\geq 0} \times \mathcal{K}_a\}^{Q_a} \times \{\mathbb{R}_{\geq 0} \times \mathcal{K}_o\}^{Q_o}$.

We represent the unavailable past information by an identifier value $r_{k_i} = (\infty, \infty)$, similar to Liu and Hauskrecht [2019]. To set covariance values concerning the unavailable information to 0, an indicator function $\mathbb{1}_k : \mathbb{R}_{\geq 0} \times \mathcal{K}_k \to \{0, 1\}$ is defined:

$$\mathbb{1}_k[r_{k_i}(\tau)] = \begin{cases} 1, & \text{if } r_{k_i,m} < \infty, r_{k_i,t} < \infty \\ 0, & \text{otherwise} \end{cases}$$

Using the indicator function, the kernel function is defined as follows:

$$K(v, v') = K_b(\tau, \tau') + \sum_{k \in \{a,o\}} \sum_{i=1}^{Q_k} \mathbb{1}[r_{k_i}(\tau)] \mathbb{1}[r_{k_i}(\tau')] K_{k_i}(r_{k_i}(\tau), r_{k_i}(\tau')),$$

where K_b is an SE kernel, input vectors $v, v' \in \mathcal{X}$ have corresponding times τ and τ', and the single-tuple kernel function $k_{k_i} : \mathbb{R}_{\geq 0} \times \mathcal{K}_{k_i} \to \mathbb{R}$ is:

$$\begin{aligned} & k_{k_i}(r_{k_i}(\tau), r_{k_i}(\tau'))) \\ &= \gamma_i \underbrace{\exp\left(-\frac{(r_{k_i,m}(\tau) - r_{k_i,m}(\tau'))^2}{\sigma_{i,m}}\right)}_{\text{Mark Kernel } K_m} \underbrace{\exp\left(-\frac{(\Delta_{k_i,t}(\tau) - \Delta_{k_i,t}(\tau'))^2}{\sigma_{i,t}}\right)}_{\text{Relative Time Kernel } K_t} \\ &= \gamma_i k_m(r_{k_i,m}(\tau), r_{k_i,m}(\tau')) k_t(\Delta_{k_i,t}(\tau), \Delta_{k_i,t}(\tau')) \end{aligned}$$

The overall kernel is a sum of the baseline SE kernel and $Q_a + Q_o$ two-dimensional single-tuple kernels K_z, each of which is an SE kernel on two dimensions: mark and relative time.

F.3 Likelihood

The likelihood for a set of treatments $\mathcal{D} = \{(t_n, m_n)\}_{n=1}^N$, can be written in terms of $\lambda^*(t, m) = \lambda^*(t) p^*(m \mid t)$:

$$p(\mathcal{D} \mid \lambda^*(\cdot)) = = \prod_n \lambda^*(t_n) p^*(m_n \mid t_n) \times \exp\Big\{ - \int_{\mathcal{T}} \lambda^*(\tau) d\tau \Big\} \tag{8}$$

The mark probability $p^*(m_n \mid t_n)$ factorizes and can be modeled by a process independent from the time intensity $\lambda^*(t)$. Then, the inference for the mark intensity $p^*(m_n \mid t_n)$ is straightforward. In the following, we derive the inference objective for the time intensity $\lambda^*(t)$.

F.3.1 Multiple Observations

We might have multiple observations $\mathcal{D} = \{\mathcal{D}_1, ..., \mathcal{D}_O\}_{o=1}^O$. For example, we can monitor a patient's blood glucose for O days and consider each day as conditionally independent given the treatment intensity. The likelihood factorizes over multiple observations:

$$p(\mathcal{D}_1, ..., \mathcal{D}_O \mid \lambda^*(\cdot)) = \prod_o p(\mathcal{D}_o \mid \lambda^*(\cdot)).$$

where the term $p(\mathcal{D}_o \mid \lambda^*(\cdot))$ follows Equation 8.

F.4 Variational Inference

In variational inference, a lower bound (ELBO), denoted by $\mathcal{L}$, of the log likelihood $\log p(\mathcal{D})$ is maximized, which is equivalent to minimizing the KL divergence between the variational distribution $q(f)$ and the true posterior $p(f \mid \mathcal{D})$.

$$\log p(\mathcal{D}) = \mathcal{L} + \mathrm{KL}[q(f) \mid\mid p(f \mid \mathcal{D})],$$

where the lower bound $\mathcal{L}$ is given by:

$$\begin{aligned}\mathcal{L} &=_q [\log p(\mathcal{D} \mid f)] - \mathrm{KL}[q(f) \mid\mid p(f)] \\ &= \mathcal{L}_{\mathcal{D}} - \mathrm{KL}[q(f) \mid\mid p(f)].\end{aligned}$$

F.4.1 Inducing Points

Inducing point approximations are commonly used for scalable inference in GPs. Let $\mathbf{Z} = \{z_m\}_{m=1}^M$ be a set of inducing points and their function evaluations are collected in a set of inducing variables $\mathbf{u} = f(Z) \in \mathbb{R}^M$, where each $u_m = f(z_m)$. Conditioned on inducing variables $\mathbf{u}$, we assume that the variational conditional distribution $q(\mathbf{f} \mid \mathbf{u})$ is equal to the true conditional $p(\mathbf{f} \mid \mathbf{u})$. Then, the variational distribution $q(\mathbf{f})$ can be written as follows:

$$q(\mathbf{f}, \mathbf{u}) = p(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u}),$$

where $q(\mathbf{u}) \sim N(\mathbf{u}; \mathbf{m}, \mathbf{S})$. Generally, we integrate out $\mathbf{u}$:

$$\begin{aligned}q(\mathbf{f}) &= \int_{\mathcal{U}} p(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u}) d\mathbf{u} = \mathcal{GP}(f; \tilde{\mu}, \tilde{\Sigma}) \\ \tilde{\mu}(\mathbf{x}) &= \mathbf{k}_{\mathbf{xz}} \mathbf{K}_{\mathbf{zz}}^{-1} \mathbf{m} \\ \tilde{\Sigma}(\mathbf{x}, \mathbf{x}') &= \mathbf{K}_{\mathbf{xx}'} - \mathbf{k}_{\mathbf{xz}} \mathbf{K}_{\mathbf{zz}}^{-1} \mathbf{k}_{\mathbf{zx}} + \mathbf{k}_{\mathbf{xz}} \mathbf{K}_{\mathbf{zz}}^{-1} \mathbf{S} \mathbf{K}_{\mathbf{zz}}^{-1} \mathbf{k}_{\mathbf{zx}'}.\end{aligned}$$

F.4.2 Evidence Lower Bound (ELBO)

The lower bound $\mathcal{L}$ is:

$$\mathcal{L} = \mathcal{L}_{\mathcal{D}} - \mathrm{KL}[q(f) \mid\mid p(f)].$$

Similar to [Lloyd et al., 2015, Matthews et al., 2016, John and Hensman, 2018], we compute the KL term $\mathrm{KL}[q(f) \mid\mid p(f)]$, by computing the KL divergence at inducing points $\mathrm{KL}[q(u) \mid\mid p(u)]$.

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_{\mathcal{D}} - \mathrm{KL}[q(\mathbf{u}) \mid\mid p(\mathbf{u})] \\ \mathcal{L}_{\mathcal{D}} &= \sum_o \sum_n {}_q[\log \lambda^*(t_n^{(o)})] - \sum_o {}_q\Big[\int_{\mathcal{T}} \lambda^*(t) dt\Big] \\ &= \sum_o \mathcal{L}_n^{(o)} - \sum_o \mathcal{L}_t^{(o)}\end{aligned}$$

First, we start with the data term $\mathcal{L}_n^{(o)}$ for a single observed point process realization, indexed by (o):

$$\begin{aligned}\mathcal{L}_n^{(o)} &= \sum_n {}_q[\log \lambda^*(t_n)] \\ &= \sum_n {}_q[\log(f(r(t_n)) + \beta)^2],\end{aligned}$$

As in [John and Hensman, 2018], we use a change of variables $f = f_n + \beta$ and exploit the computational trick used in [Lloyd et al., 2015] to compute the one-dimensional integral ${}_q[\log f_n^2]$:

$$\begin{aligned}{}_q[\log f^2] &= \int_{-\infty}^{\infty} \log(f^2) \mathcal{N}(f; \tilde{\mu}, \tilde{\sigma}^2) df \\ &= -\tilde{G}\Big(-\frac{\tilde{\mu}^2}{2\tilde{\sigma}^2}\Big) + \log\Big(\frac{\tilde{\sigma}^2}{2}\Big) - C\end{aligned}$$

where C is the Euler-Mascheroni constant and $\tilde{G}$ is defined by the confluent hyper-geometric function, which is approximated by a look-up table for faster computation.

To compute the integral term $\mathcal{L}_t$, we divide the observation period into $N+1$ intervals with end points $(0, t_1, \ldots, t_N, T)$, where $\{t_i\}_{i=1}^N$ includes all event points that have an affect on the conditional intensity value. In each interval $[t_{n-1}, t_n]$, the conditional intensity values $\{\lambda^*(\tau) : \tau \in [t_{n-1}, t_n]\}$ are estimable using the past history $\mathcal{H}_{<t_n}$. Using a shorthand notation for $r(t) = \mathbf{r} \in \mathcal{X}$, we write the integral term as a sum of three integrals:

$$
\begin{aligned}
\mathcal{L}_t &= \sum_{o=1}^{O} {}_q\Big[\int_{\mathcal{T}} \lambda^*(\tau) d\tau\Big] \\
&= \sum_{o=1}^{O} {}_q\Big[\int_{\mathcal{T}} (f(r(\tau)) + \beta)^2 d\tau\Big] \\
&= \sum_{o=1}^{O} \Big[\int_{\mathcal{T}} \Big({}_q[f^2(\mathbf{r})] + 2\beta\, {}_q[f(\mathbf{r})] + \beta^2\Big) d\tau\Big] \\
&= \sum_{o=1}^{O} \Big[\int_{\mathcal{T}} {}_q[f(\mathbf{r})]^2 d\tau + \int_{\mathcal{T}} \mathrm{Var}_q[f(\mathbf{r})] d\tau + 2\beta \int_{\mathcal{T}} {}_q[f(\mathbf{r})] d\tau + \beta^2 |\mathcal{T}_o|\Big] \\
&= \sum_{o=1}^{O} \sum_{n=1}^{N_o+1} \Big[\int_{t_{n-1}}^{t_n} {}_q[f(\mathbf{r})]^2 d\tau + \int_{t_{n-1}}^{t_n} \mathrm{Var}_q[f(\mathbf{r})] d\tau + 2\beta \int_{t_{n-1}}^{t_n} {}_q[f(\mathbf{r})] d\tau + \beta^2 |\mathcal{T}_o|\Big].
\end{aligned}
$$

We can compute each integral term as follows:

$$
\begin{aligned}
\int_{t_{n-1}}^{t_n} {}_q[f(r(\tau))]^2 d\tau &= \mathbf{m}^{\mathbf{T}} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{\Psi}_{\mathbf{n}} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{m} \\
\int_{t_{n-1}}^{t_n} \mathrm{Var}_q[f(\tau)] d\tau &= \sum_{d=1}^{D} \gamma_d \int_{t_{n-1}}^{t_n} \mathbb{1}[h_d(t)] dt + \mathrm{Tr}(\mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{\Psi}_{\mathbf{n}}) + \mathrm{Tr}(\mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{S} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{\Psi}_{\mathbf{n}}) \\
\int_{t_{n-1}}^{t_n} {}_q[f(\tau)] d\tau &= \mathbf{\Phi}_{\mathbf{n}}^{\mathbf{T}} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{m} \\
\mathbf{\Psi}_{\mathbf{n}}(\mathbf{z}, \mathbf{z}') &= \int_{t_{n-1}}^{t_n} K(\mathbf{z}, r(\tau)) K(r(\tau), \mathbf{z}') d\tau \\
\mathbf{\Phi}_{\mathbf{n}}(\mathbf{z}) &= \int_{t_{n-1}}^{t_n} K(\mathbf{z}, r(\tau)) d\tau.
\end{aligned}
$$

Then, all integral terms are plugged in:

$$
\begin{aligned}
\mathcal{L}_t^{(o)} &= \mathbf{m}^{\mathbf{T}} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{\Psi}_{\mathbf{n}} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{m} \\
&+ \sum_{d=1}^{D} \gamma_d \int_{t_{n-1}}^{t_n} \mathbb{1}[h_d(t)] dt + \mathrm{Tr}(\mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{\Psi}_{\mathbf{n}}) + \mathrm{Tr}(\mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{S} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{\Psi}_{\mathbf{n}}) \\
&+ 2\beta \mathbf{\Phi}_{\mathbf{n}}^{\mathbf{T}} \mathbf{K}_{\mathbf{zz}}^{-\mathbf{1}} \mathbf{m} + \beta^2 |\mathcal{T}_o|.
\end{aligned}
$$

The phi vector $\mathbf{\Phi}_{\mathbf{n}}$ and the psi matrix $\mathbf{\Psi}_{\mathbf{n}}$ have closed form solutions, obtained by evaluating the integrals for the sum of SE kernels. The only difference to Liu and Hauskrecht [2019] is the inclusion of the SE mark kernel terms $k_m(\cdot, \cdot)$.

$$
\begin{aligned}
\Phi_n(z) = \sum_{q=1}^{Q_a+Q_o} \mathbb{1}[z_q] \mathbb{1}[r_q(t_n)] \gamma_q \frac{\sqrt{\pi \alpha_{q,t}}}{\sqrt{2}} \exp\left(\frac{(m_{z_q} - r_{q,m}(t_n))^2}{\alpha_{q,m}}\right) \\
\left[\mathrm{erf}\left(\frac{t_n - r_{q,t}(t_n) - t_{z_q}}{\sqrt{2\alpha_{q,t}}}\right) - \mathrm{erf}\left(\frac{t_{n-1} - r_{q,t}(t_n) - t_{z_q}}{\sqrt{2\alpha_{q,t}}}\right)\right],
\end{aligned}
$$

$$
\begin{aligned}
\Psi_n(z, z') = \sum_{i,j=1}^{Q_a+Q_o} & \mathbb{1}[z_i]\mathbb{1}[z_j]\mathbb{1}[r_i(t_n)]\mathbb{1}[r_j(t_n)]\gamma_i\gamma_j \frac{\sqrt{\pi\alpha_{i,t}\alpha_{j,t}}}{\sqrt{2(\alpha_{i,t}+\alpha_{j,t})}} \\
& \exp\left(\frac{(m_{z_i} - r_{i,m}(t_n))^2}{\alpha_{i,m}}\right) \exp\left(\frac{(m_{z_j} - r_{j,m}(t_n))^2}{\alpha_{j,m}}\right) \\
& \exp\left(-\frac{(t_{z_i} + r_{i,t}(t_n) - t'_{z'_j} - r_{j,t}(t_n))^2}{2(\alpha_{i,t}+\alpha_{j,t})}\right) \\
& \left[\operatorname{erf}\left(\frac{\alpha_i(t_n - r_{j,t}(t_n) - t'_{z'_j}) + \alpha_j(t_n - r_{i,t}(t_n) - t_{z_i})}{\sqrt{2\alpha_{i,t}\alpha_{j,t}(\alpha_{i,t}+\alpha_{j,t})}}\right)\right. \\
& \left. - \operatorname{erf}\left(\frac{\alpha_i(t_{n-1} - r_{j,t}(t_n) - t'_{z'_j}) + \alpha_j(t_{n-1} - r_{i,t}(t_n) - t_{z_i})}{\sqrt{2\alpha_{i,t}\alpha_{j,t}(\alpha_{i,t}+\alpha_{j,t})}}\right)\right].
\end{aligned}
$$

G Outcome Model Details

We model the outcome trajectory $\mathbf{Y} = \{y(\tau) : \tau \in \mathbb{R}_{\geq 0}\}$ by a conditional GP model, combining three independent function components: (i) a baseline progression, (ii) treatment effects and (iii) a noise variable [Schulam and Saria, 2017, Zhang et al., 2020]:

$$
y(\tau) = \underbrace{f_b(\tau)}_{\text{Baseline}} + \underbrace{f_a(\tau; \mathbf{a})}_{\text{Treatment Effect}} + \underbrace{\epsilon(\tau)}_{\text{Noise}}, \tag{9}
$$

Baseline progression. The baseline progression $f_b(\tau)$ over time τ can be modeled by a GP prior. The kernel function is chosen, depending on the application context. For example, a baseline function $f_b(\tau) \sim \mathcal{GP}(0, K_{se} + K_{per})$, with a zero mean and a kernel function equal to the sum of a squared exponential (SE) and a periodic (PER) kernel was proposed to model blood pressure and the heart rate measurements in [Cheng et al., 2020]. A combination of two non-stationary kernel functions was chosen as the baseline kernel function to model creatinine measurements in [Schulam and Saria, 2017]. In [Xu et al., 2016], the baseline kernel function was chosen as the sum of a linear and an exponential kernel. In our experiments, we choose a combination of a linear and an SE kernel to model the baseline progression of blood glucose measurements, as non-diabetic patients mostly have linear or constant baseline progressions, with a significant amount of short-term noise [Zhang et al., 2020, Ashrafi et al., 2021].

Treatment Effect Model. We define the treatment effect function $f_a(\tau; \mathbf{a})$ as a Gaussian process: $f_a(\tau; \mathbf{a}) \sim \mathcal{GP}$, similar to Cheng et al. [2020]. While a latent force model (LFM) [Alvarez et al., 2009] extension is proposed in Cheng et al. [2020], we use an SE kernel, as it produces sufficient performance for our use case. For each treatment, we consider the treatment effect as additive:

$$
f_a(\tau; \mathbf{a}) = \sum_{a_i = (t_i, m_i) \in \mathbf{a}} f_m(m_i) f_t(\tau; t_i),
$$

where the scaling function $f_m(m_i) : \mathcal{K}_a \to \mathbb{R}$ is a function of the treatment mark $m_i \in \mathcal{K}_a$ and $\mathcal{K}_a$ denotes the treatment mark space. Each treatment type is assumed to have a distinct pair of a scaling function f_m and a time-dependent effect function f_t. We assume the scaling function f_m is shared among treatments of the same individual and the time-dependent effect function f_t is shared among treatments of the same treatment type.

The scaling function $f_m^{(r)}(m_i)$ is a linear function that captures the individualized treatment response magnitude for each patient (r):

$$
f_m^{(r)}(m_i) = \beta_0^{(r)} + \beta_1^{(r)} m_i,
$$

where a hierarchical Gaussian prior is placed on parameters $\{\beta_0^{(r)}, \beta_1^{(r)}\}$: $\beta_0^{(r)} \sim \mathcal{N}(\beta_0, \sigma_0^2)$, $\beta_1^{(r)} \sim \mathcal{N}(\beta_1, \sigma_1^2)$.

The time-dependent function $f_t(\tau; \mathbf{a})$ is defined as a piecewise function that is equal to 0 outside treatment effect intervals $\mathcal{T}_{t_i} = [t_i, t_i + T_i]$ and $\mathcal{T}_{t'_i} = [t'_i, t'_i + T'_i]$:

$$f_t(\tau; t_i) = \begin{cases} \mathcal{GP}(0, k_{f_t, f'_t}(\tau, \tau'; t_i, t'_i)), & \text{if } \tau \in \mathcal{T}_{t_i}, \tau' \in \mathcal{T}_{t'_i} \\ 0, & \text{otherwise} \end{cases},$$

where the treatment effect interval T_i is set by the domain knowledge. The kernel function works with non-negative relative times $\Delta\tau_i = \text{RELU}(\tau - t_i)$, rather than the absolute time τ:

$$k_{f_t, f'_t}(\tau, \tau'; t_i, t'_i) = \exp\left(-\frac{(\Delta\tau_i - \Delta\tau'_i))^2}{l_t^2}\right).$$

The adoption of the piecewise definition and relative times is enforcing a 'causal' kernel for the treatment effect, as in 'causal' GP proposed in Cunningham et al. [2012]. Here, the term 'causal' is used differently than the causality discussion in previous sections and has more like a common sense meaning implying the time direction.

H Learning

As the joint likelihood factorizes in time-order, we can maximize treatment and outcome model parameters independently, conditioned on the history. For the treatment model, we maximize the ELBO $\mathcal{L}$ with respect to variational parameters $\theta_{a,z} = \{\mathbf{m_z}, \mathbf{S_z}\}$ and kernel hyperparameters $\theta_{a,h} = \{\gamma_q, \alpha_q\}_{q=0}^{Q_a+Q_o}$. For the outcome model, we maximize the marginal likelihood $p(\mathcal{D} \mid \theta)$ with respect to model hyperparameters $\theta_o = \{\ell_t, \beta_0, \beta_1, \{\beta_0^{(r)}, \beta_1^{(r)}\}_{r=1}^{N_{patient}}\} \cup \theta_{o,b}$, where $\theta_{o,b}$ denote the parameters of the baseline function whose definition depends on the application context and $\{\beta_0, \beta_1, \beta_0^{(r)}, \beta_1^{(r)}\}$ denote the parameters of the hierarchical linear model for the scale function $f_m^{(r)}(\cdot)$.

I Inference

I.1 Test Log-Likelihood (TLL)

The outcome model is a GP prior with an independent Gaussian noise. The posterior of the latent function $\mathbf{f_q}$ at query times $\mathbf{q}$ follows a multivariate normal distribution and its mean and covariance functions have a closed form solution: $\mathbf{f_q} \mid \mathbf{y} \sim \mathcal{N}(\tilde{\mu}_q, \tilde{\Sigma}_q)$, with $\hat{\mu}_q = \mathbf{k}_*^{\mathbf{T}}(\mathbf{K} + \sigma_\epsilon^2\mathbf{I})^{-1}\mathbf{y}$ and $\hat{\Sigma}_q = \mathbf{K}_{**} - \mathbf{k}_*^{\mathbf{T}}(\mathbf{K} + \sigma_\epsilon^2\mathbf{I})^{-1}\mathbf{k}_*$. The test log-likelihood has a closed form solution and can be computed by integrating out the posterior $\mathbf{f_q} \mid \mathbf{y}$:

$$\log p(\mathcal{D}_{test} \mid \mathcal{D}) = \log \int_{\mathbf{f_q}} \log p(\mathcal{D}_{test} \mid \mathbf{f_q}) p(\mathbf{f_q} \mid \mathcal{D}) d\mathbf{f_q}.$$

For the treatment model, the marginal log-likelihood is intractable. A lower-bound on the test log-likelihood $\log p(\mathcal{D}_{test} \mid \mathcal{D})$ can be computed as follows:

$$\begin{aligned} \log p(\mathcal{D}_{test} \mid \mathcal{D}) &= \log \int_{g^*} \log p(\mathcal{D}_{test} \mid g^*) p(g^* \mid \mathcal{D}) dg^* \\ &= \log_{p(g^* \mid \mathcal{D})} \left[p(\mathcal{D}_{test} \mid g^*)\right] \\ &\approx \log_{q(g^*)} \left[p(\mathcal{D}_{test} \mid g^*)\right] \\ &\geq_{q(g^*)} \left[\log p(\mathcal{D}_{test} \mid g^*)\right], \end{aligned}$$

where we approximate the true posterior $p(g^* \mid \mathcal{D})$ by the variational distribution $q(g^*) = p(g^* \mid u)q(u)$. The term ${}_{q(g^*)}\left[\log p(\mathcal{D}_{test} \mid g^*)\right]$ can be computed as the steps detailed in Appendix F.4.2, as it is in the same form with the expected log likelihood component of the ELBO.

I.2 Counterfactual Query

For a policy counterfactual, we estimate the potential outcome of the past trajectory $P(\mathbf{Y}_{\leq T}[\tilde{\pi}_{\leq T}] \mid \mathcal{H}_{\leq T}, \pi_{\leq T})$, under a new treatment policy specified by $\tilde{\pi}_{\leq T}$ different than the observed policy

specified by $\pi_{\leq T}$. To answer a counterfactual query, we require access to the posterior distribution of noise variables in addition to the interventional intensity λ^*_{int}, which we denote by λ^*_{cf} to emphasize the distribution we are interested in.

For the outcome model, the posterior of the noise variables is equal to the difference between the posterior of the latent function $f(\tau) \mid \mathbf{y}$ and the observed outcome $y(\tau)$:

$$\epsilon(\tau) \mid \mathbf{y} = y(\tau) - f(\tau) \mid \mathbf{y}.$$

We use a point estimate for the noise variable $\epsilon(\tau)$ for each time τ:

$$\big[\epsilon(\tau) \mid \mathbf{y}\big] = y(\tau) - \big[f(\tau) \mid \mathbf{y}\big].$$

Regarding the noise variables of the treatment model, we extend the counterfactual sampling algorithm for non-homogeneous Poisson processes provided in [Noorbakhsh and Rodriguez, 2021] to history-dependent point processes where future events depend on past events. The key difference in algorithms is that the counterfactual sampling algorithm provided in [Noorbakhsh and Rodriguez, 2021] is based on Lewis' thinning algorithm [Lewis and Shedler, 1979], while our algorithm is based on Ogata's thinning algorithm [Ogata, 1981], which performs thinning in small intervals. The thinning procedure first samples points from an upper-bounding intensity and then performs rejection sampling with a probability proportional to the target intensity, so that the intensity of accepted points is equivalent to the target intensity.

I.2.1 Ogata's Thinning Algorithm

Consider an initial time point τ, where we start sampling. Ogata's algorithm starts with specifying an interval such as $[\tau, \tau + s(\tau)]$, where the function $l(\cdot)$ specifies the end point $\tau + l(\tau)$ until which it is possible to upper bound the conditional intensity λ^* [Rasmussen, 2011]. The upper bound intensity should be greater than or equal to the supremum of the conditional intensity λ^* in the interval $[\tau, \tau + l(\tau)]$:

$$\lambda_{ub} \geq \sup_{s \in [\tau, \tau + l(\tau)]} \lambda^*(s).$$

Using the constant upper bound intensity, one can sample a candidate point from a Poisson process $t_i \sim \mathcal{PP}(\lambda_{ub})$. After sampling t_i, the algorithm considers two options:

- If $t_i > l(\tau)$, then there exists no points in the interval $[\tau, \tau + l(\tau)]$ and the sampling procedure repeats with the next time interval.
- If $t_i < l(\tau)$, t_i is kept with the probability $\lambda^*(t_i)/\lambda_{ub}$, which produces the target intensity by re-weighting the upper bound with the factor $\lambda^*(t_i)/\lambda_{ub}$. In practice, a noise variable U_i can be sampled from a uniform distribution: $u_i \sim \mathcal{U}(0, \lambda_{ub})$ and the point can be kept if $u_i \leq \lambda^*(t_i)$. Regardless of the thinning decision, the procedure repeats from the time point t_i.

The sampling process is illustrated in Algorithm 3.

[H]

Figure 3: Ogata's Thinning algorithm

Input: Start T_1, End T_2, Interval function $l(\cdot)$, Conditional intensity $\lambda^*(\cdot)$
Output: Point process sample $\mathcal{T} = \{t_1, \ldots, t_n\}$ [1] SampleT_1, T_2, l, λ^* $\tau = T_1, \mathcal{T} = \emptyset$. Initialize $\tau < T_2$ $\lambda_{ub} = \sup_{s \in [\tau, \tau + l(\tau)]} \lambda^*(s)$. Upper-bound. $t_i \sim Exp(\lambda_{ub})$. Draw the inter-arrival time. $t_i \leq l(\tau)$ and $t_i + \tau \leq T$ Candidate in interval. $u_i \sim \mathcal{U}(0, \lambda_{ub})$. Draw noise variable. $u_i \leq \lambda^*(\tau + t_i)$ Keep with probability λ^*/λ_{ub}. $\mathcal{T} = \mathcal{T} \cup \{\tau + t_i\}$. No point in interval. $\tau = \tau + l(\tau)$. $\mathcal{T} = \{t_1, \ldots, t_n\}$.

I.2.2 Counterfactual Sampling Algorithm for TPPs

Our algorithm assumes the data generative process follows Ogata's thinning algorithm [Ogata, 1981], where thinned events and noise variables are considered as latent variables. For a given interval $[\tau, \tau + l(\tau)]$, the upper-bound intensity λ_{ub} needs to be an upper bound for both the observational

intensity λ^*_{obs} and the counterfactual intensity λ^*_{cf}, so that accepted events, thinned events and the noise posteriors are well defined:

$$\lambda_{ub} \geq \sup_{\tau \in \mathcal{I}} \{\lambda^*(\tau) : \lambda^* \in \{\lambda^*_{obs}, \lambda^*_{cf}\}\}.$$

In practice, we select the function $l(\tau)$ so that it returns the next event after time τ, which can be a future treatment or an outcome. As a result, the observation period $[0, T]$ is split into intervals with end points $(0, t_1, \ldots, t_N, T)$, where in each interval, it is possible to compute both the observational intensity λ^*_{obs} and the counterfactual intensity λ^*_{cf}, conditioned on their respective histories $\mathcal{H}_{obs}$ and $\mathcal{H}_{cf}$. Here, t_N denotes the time of the last event, among all event times $\mathbf{t_a} \cup \mathbf{t_o}$.

As the observed data does not contain thinned events, our algorithm first samples thinned events by using the difference of the upper bound intensity and the observational intensity λ_{obs}: $\lambda^*_{rej} = \lambda_{ub} - \lambda^*_{obs}$, in the interval $[\tau, \tau + l(\tau)]$. Second, it computes the posterior distribution of uniform noise variables U_i for the next point, e.g. a rejected point t_{rej} or an observed point t_{obs}. By inspecting the sampling process in Algorithm 3, we can compute the noise posterior for a noise variable U_i as follows:

$$p(U_i \mid t_i, \lambda_{ub}, \lambda^*_{obs}) = \begin{cases} \mathcal{U}(0, \lambda^*_{obs}(t_i)), & \text{if } t_i \text{ is observed} \\ \mathcal{U}(\lambda^*_{obs}(t_i), \lambda_{ub}), & \text{if } t_i \text{ is rejected.} \end{cases}$$

Having access to (i) observed (accepted) events, (ii) rejected events, (iii) posterior distributions of noise variables, (iv) the observational intensity and (v) the counterfactual intensity, we can sample point process realizations from the counterfactual distribution, in a similar fashion to the forward sampling process of Ogata's algorithm. The counterfactual sampling process is presented in Algorithm 4. There are three key differences to the observational sampling process in Algorithm 3: (1) in each interval, the algorithm samples observational rejected events using the intensity $\lambda_{ub} - \lambda^*_{obs}$, (2) the counterfactual algorithm uses the noise posteriors instead of noise priors $\mathcal{U}(0, \lambda_{ub})$ and (3) the accept/reject decision are based on the counterfactual intensity λ^*_{cf} instead of the observational intensity λ^*_{obs}.

[H]

Figure 4: Counterfactual Sampling Algorithm for TPPs (Based on Ogata's)

Input: Start T_1, End T_2, Interval function $l(\cdot)$, Observational intensity $\lambda^*_{obs}(\cdot)$, Counterfactual intensity $\lambda^*_{cf}(\cdot)$, Observed points $\mathcal{T}_{obs}$

Output: Counterfactual point process sample $\mathcal{T}_{cf} = \{t_1, \ldots, t_n\}$ [1] SampleT, l, λ^* $\tau = T_1, \mathcal{T} = \emptyset$. Initialize $\tau < T_2$ $\lambda_{ub} = \sup_{s \in [\tau, \tau + l(\tau)]} \{\lambda^*(s) : \lambda^* \in \{\lambda^*_{obs}, \lambda^*_{cf}\}\}$. Upper-bound. $t_{rej} = \textsc{Sample}(\tau, \tau + l(t), \lambda_{ub}, \lambda_{ub} - \lambda^*_{obs})$. Draw a rejected point. $t_{rej} \leq l(\tau)$ and $t_{rej} + \tau \leq T$ $u_{rej} \sim \mathcal{U}(\lambda^*_{obs}(\tau + t_r, \mathcal{H}_{obs}), \lambda_{ub})$. Noise posterior. $u_{rej} \leq \lambda^*_{cf}(\tau + t_r)$ $\mathcal{T}_{cf} = \mathcal{T}_{cf} \cup \{\tau + t_r\}$. $\tau = \tau + t_r$. No rejections in interval. $\tau + l(\tau) \in \mathcal{T}_{obs}$ Check if the end point is observed. $t_{obs} = \tau + l(\tau)$. $u_{obs} \sim \mathcal{U}(0, \lambda^*_{obs}(t_{obs}, \mathcal{H}_{obs}))$. Noise posterior. $u_{obs} \leq \lambda^*_{cf}(t_{obs}, \mathcal{H}_{cf})$ $\mathcal{T}_{cf} = \mathcal{T}_{cf} \cup \{t_{obs}\}$. $\tau = \tau + l(\tau)$. $\mathcal{T}_{cf} = \{t_1, \ldots, t_n\}$.

J Experiment Details

J.1 Real-World Data Experiment Details

J.1.1 Data Preprocessing

Our target real-world data set consists of meal-blood glucose measurements for 14 non-diabetic individuals, where individuals are monitored over a 3-day period [Zhang et al., 2020]. The blood glucose data is measured via a continuous monitoring device, which approximately takes a sample every 15 minutes. The meal data is collected through a meal diary, recorded daily by individuals. Later, meal records in the diary are translated into nutrient values by a look-up table.

As the meal data set is collected by individuals, it inherently contains measurement errors [Zhang et al., 2020]. To deal with the measurement errors, we take a data preprocessing step before fitting our treatment model to the data set. For our application, there are two main measurement errors that

may lead to incorrect results: (i) individuals may record their meals in parts, where a single meal is recorded as 2-3 meal occurrences, and (ii) individuals record their meal times with error.

To deal with the errors of type (i), we remove redundant meals by keeping only a single meal event if two meals co-occur in a 2-hour interval. To deal with the errors of type (ii), we update the meal time if a meal seems to occur at a time where the blood glucose is changing rapidly. For example, if an individual reports a meal with a delay, the meal seems to occur at a time point where the blood glucose measurements have already increased rapidly and is about to decline. The recorded meal time (with error) implies an incorrect result that the meal leads to a decrease in the blood glucose, or similarly, that the meal likelihood increases with increasing blood glucose. On the contrary, the bump in the glucose level is most likely caused by the same meal in the first place. In practice, we move a meal time to the next possible time point where the blood glucose derivative is less than a threshold, 0.5. At each time point t_i, we approximate the derivative of the blood glucose by averaging derivative values for two consecutive regions, (t_{i-1}, t_i) and (t_i, t_{i+1}):

$$dy_i \approx \frac{1}{2}\left(\frac{(y_i - y_{i-1})}{t_i - t_{i-1}} + \frac{(y_{i+1} - y_i)}{t_{i+1} - t_i}\right).$$

J.1.2 Model Definition

Treatment Intensity. We define five model variants for each patient, since we assume each patient's meal habits are different, corresponding to a distinct treatment policy. As a result, each treatment model is defined and trained independently on each patient's meal–glucose measurements. Each model uses one of the intensities $\{\lambda_b, \lambda^*_{ba}, \lambda^*_{bo}, \lambda^*_{ao}, \lambda^*_{bao}\}$ defined in Experiments section. The constant baseline intensity β_0 is set to 0.1 for all experiments.

We choose model parameters $\{Q_a, Q_o\}$ as $Q_a = 1, Q_o = 1$, i.e., the meal intensity depends only on the last treatment and the last outcome observations. Furthermore, we assume treatment marks do not have an affect on the next meal time. Even though these assumptions lead to a simplification of complex real-world dynamics between meals and the blood glucose, our goal is not to model these dynamics in the best possible way, but rather to understand if our joint model is useful in handling time-varying confounding.

At a query time τ, the retrieval function $r(\tau)$ returns the input vector $r(\tau) = \mathbf{r} \in \mathbb{R}^4$, as defined in F.2. We choose $M = 20$ inducing points: $\mathbf{Z} \in \mathbb{R}^{M\times 4}$. We place the inducing points in regular intervals inside a target region, specified by the treatment intensity component g_b, g^*_a or g^*_o. For example, the baseline component's input domain is the observation period $[0, T]$ and inducing points for the baseline dimension are placed onto $[0, T]$ regularly. Similarly, input domains of treatment-effect and outcome-effect kernels include relative times to the last Q_a and Q_o events. For these two dimensions, we place inducing points between 0 and the maximum (or some quantile) time between the last Q_a or Q_o events. Inducing points are assumed to be independent along each dimension similar to Liu and Hauskrecht [2019].

We choose kernel hyperparameters $\theta_{a,h} = \{\gamma_q, \ell_q\}_{q=0}^{Q_a+Q_o}$ by inspection, since our data set is small and there exists identifiability issues when one optimizes the effect of each component on the treatment effect simultaneously. As the number of meal events are small in the data set, we choose relatively small values for the variance parameters $(\gamma_b, \gamma_a, \gamma_o) = (0.1, 0.05, 0.15)$. We choose lengthscale parameters as $(\ell_b, \ell_a, \ell_{o,t}, \ell_{o,m}) = (7.0, 1.0, 100.0, 2.5)$. Each lengthscale value corresponds to an underlying assumption: (i) the baseline length scale $\ell_b = 7.0$ assumes a slow-changing smooth baseline function, (ii) the treatment-effect time lengthscale ℓ_a is 1.0 as the last meal generally has an effect on the next meal in the next 3-4 hours time span, (iii) a very large outcome-effect time lengthscale $\ell_{o,t} = 100$ is chosen so that the outcome-effect is constant in time for a given glucose value y, mimicking a simultaneous outcome-effect on the meal intensity with a large number of regular glucose measurements, (iv) the outcome-effect mark lengthscale $\ell_{o,m} = 2.5$ is chosen so that the mark effect is also a slow-changing, smooth function, as the glucose level should gradually affect meal intensity.

Outcome Model. We define a single hierarchical outcome model for all patients, as detailed in Appendix G.

The baseline progression for each patient (r) is the sum of a linear function and an SE kernel to capture short-term fluctuations. The linear mean function has an intercept and a slope parameter

Table 1: Results for the test log likelihood (TLL) values for treatment models with distinct meal intensities, for all patients. Models are trained on first 2-days of meal-glucose data. The last day is used for computing the TLL. The intensity $\lambda^*_{ba}(\tau)$ provide the highest TLL value for all patients.

		MEAL INTENSITIES				
PATIENT-ID	METRIC	λ^*_b	λ^*_{ba}	λ^*_{bo}	λ^*_{ao}	λ^*_{bao}
PATIENT 1	14*TLL	-12.49	$\mathbf{-5.03}$	-9.88	-7.68	-10.68
PATIENT 2		-10.55	$\mathbf{-2.50}$	-10.57	-10.57	-11.85
PATIENT 3		-13.07	$\mathbf{-5.94}$	-12.11	-12.78	-12.79
PATIENT 4		-8.39	$\mathbf{4.62}$	-9.38	-9.64	-11.51
PATIENT 5		-5.80	$\mathbf{2.49}$	-11.32	-10.96	-12.65
PATIENT 6		-11.65	$\mathbf{18.44}$	-8.53	-8.59	-9.73
PATIENT 7		-10.93	$\mathbf{6.51}$	-12.49	-10.89	-13.97
PATIENT 8		-12.24	$\mathbf{2.17}$	-10.90	-9.70	-12.20
PATIENT 9		-13.35	$\mathbf{5.41}$	-13.07	-12.92	-13.95
PATIENT 10		-14.79	$\mathbf{6.37}$	-11.42	-11.39	-12.60
PATIENT 11		-7.30	$\mathbf{0.44}$	-12.38	-11.23	-13.79
PATIENT 12		-15.71	$\mathbf{-3.18}$	-12.11	-13.60	-14.80
PATIENT 13		-4.73	$\mathbf{1.40}$	-10.09	-9.99	-11.67
PATIENT 14		-15.22	$\mathbf{4.48}$	-10.93	-8.95	-12.88
MEAN	TLL	-11.16	$\mathbf{2.56}$	-11.08	-10.63	-12.50

$\{c^{(r)}, d^{(r)}\}$, initialized to $\{1.0, 1.0\}$. Each short-term noise SE kernel has parameters $\{\gamma^{(r)}_{st}, \ell^{(r)}_{st}\}$, initialized to $\{0.1, 1.0\}$.

The time-dependent treatment function $f_a(\tau; \mathbf{a})$ has an SE kernel with hyperparameters $\{\ell^{(m)}_t\}$ for each treatment type (m). We assume all meals are of the same type and initialize the lengthscale to 0.5: $\ell^{(m)}_t = \ell_t = 0.5$. The effective interval $\mathcal{T}_{t_i}$ for each meal t_i is set to 3 hours: $\mathcal{T}_{t_i} = 3$. The scaling function $f^{(r)}_m$ that specifies the magnitude of the treatment response has an intercept and a slope parameter $\{\beta^{(r)}_0, \beta^{(r)}_1\}$ for each patient (r). A hierarchical model is imposed on the intercept and the slope parameters $\{\beta^{(r)}_0, \beta^{(r)}_1\}$ by defining a hierarchical Gaussian prior on them: $\beta^{(r)}_0 \sim \mathcal{N}(\beta_0, \sigma^2_0), \beta^{(r)}_0 \sim \mathcal{N}(\beta_1, \sigma^2_1), \forall(r) \in \{1, \ldots, 14\}$. For these parameters, we choose the following initialization: $\{\beta_0, \beta_1, \sigma_0, \sigma_1\} = \{0.1, 0.1, 0.1, 0.1\}$.

Training. We assume each day of meal-glucose observations as conditionally independent given model components. We train each model on the first two days of observations and use the third day as the test set.

For the outcome model, we have a Gaussian likelihood, i.e. an exact GP model. We maximize the log marginal likelihood with respect to the hyperparameters $\theta_o = \{\beta_0, \beta_1, \{c^{(r)}, d^{(r)}, \gamma^{(r)}_{st}, \ell^{(r)}_{st}, \beta^{(r)}_0, \beta^{(r)}_1\}^{14}_{r=1}\}$. We fix parameters $\{\sigma_0, \sigma_1\}$ at their initialized values.

For the treatment model, we maximize the training objective ELBO with respect to the variational parameters $\theta_{a,z} = \{\mathbf{m_z}, \mathbf{S_z}\}$. Kernel hyperparameters $\theta_{a,h}$ are selected according to the procedure detailed above.

Inference. For the treatment model, we report the test log-likelihood (TLL) for the third day to compare treatment model variants. TLL results for all patients are shown in Table 1. In agreement with the mean TLL value, the intensity λ^*_{ao} produce largest TLL values for all patients. In addition, either the baseline intensity λ_b with no treatment-outcome effect or the intensity λ^*_{bao} with all components provides lowest TLL values for all patients. This suggests that a treatment model independent of the past history is not able to explain the meal intensity well. Similarly, the intensity with all components seem to be overly complex for this small data set.

J.2 Simulation Experiment Details

We perform a simulation study to evaluate our model's performance on the causal tasks. Our joint model is composed of a treatment and an outcome model. For the treatment model, we choose a treatment- and outcome-dependent intensity λ^*_{ao} to include both sources of time-varying confounding. For the outcome model, we choose the hierarchical outcome model, same as the real-world data experiment. The model definition and the training procedure are detailed in Appendix J.1.2.

We fit our joint model to the real-world meal–glucose data set to obtain ground-truth model components $\mathcal{M}^{(r)}_{gt} = \{\lambda^*_{\pi^{(r)}}(\tau), f^{(r)}_b(\tau), f^{(r)}_a(\tau; \mathbf{a})\}$ for each patient $(r) \in \{1, \ldots, 14\}$. We use the learned ground-truth components $\{\mathcal{M}^{(r)}_{gt}\}^{14}_{r=1}$ to simulate hypothetical patients.

We divide simulated patients into two policy groups $\{\pi_A, \pi_B\}$, representing distinct treatment policies of different hospitals, different countries, etc. Treatment intensities of two policies correspond to learned intensity functions of two real-world patients: Patients {4,12}. We choose these two intensities, so that the distributions of sequential treatments induced by two policies show different characteristics. For example, Patient 4 has 5 meals on average per day, while Patient 12 has 3 meals on average per day, which leads to higher intensity values for Patient 4 compared to the intensity of Patient 12. The ground-truth functions g^*_a and g^*_o, that are learned from Patients {4,12} are shown in Figure 5.

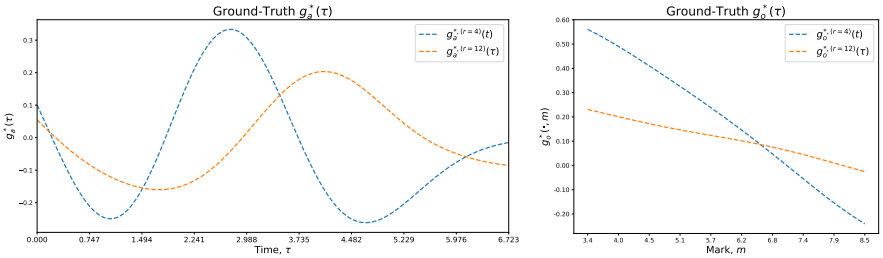

Figure 5: The learned ground-truth functions g^*_a and g^*_o for Patient 4 and Patient 12. (Left) After a meal occurs, both intensities first decrease and then increase to account for the next meal. The bump for the next meal occurs earlier for Patient 4 (blue dashed line) than Patient 12 (orange dashed line). Additionally, the magnitude of the bump is higher for Patient 4 than Patient 12. (Right) The meal intensity of Patient 4 (blue dashed line) increases more than Patient 12 (orange dashed line) as the blood glucose decreases.

To enable individualization among the blood glucose dynamics of patients, we assume there are three patient groups $\{gr_1, gr_2, gr_3\}$. Patients in the same group share the baseline function $f^{(r)}_b(\tau)$ and the treatment response function $f^{(r)}_a(\tau; \mathbf{a})$, e.g. for patients in the group gr_1 have the following functions: $\{f_{b,gr_1}, f_{a,gr_1}\}$.

We choose three ground-truth functions corresponding to the learned outcome models of three real-world patients (Patients {2,3,8}), by qualitatively assessing the learned functions with respect to the domain knowledge and findings about the effect of a meal on the blood glucose in Wyatt et al. [2021]. The train and test fits of the outcome models for these patients are shown in Figure 6.

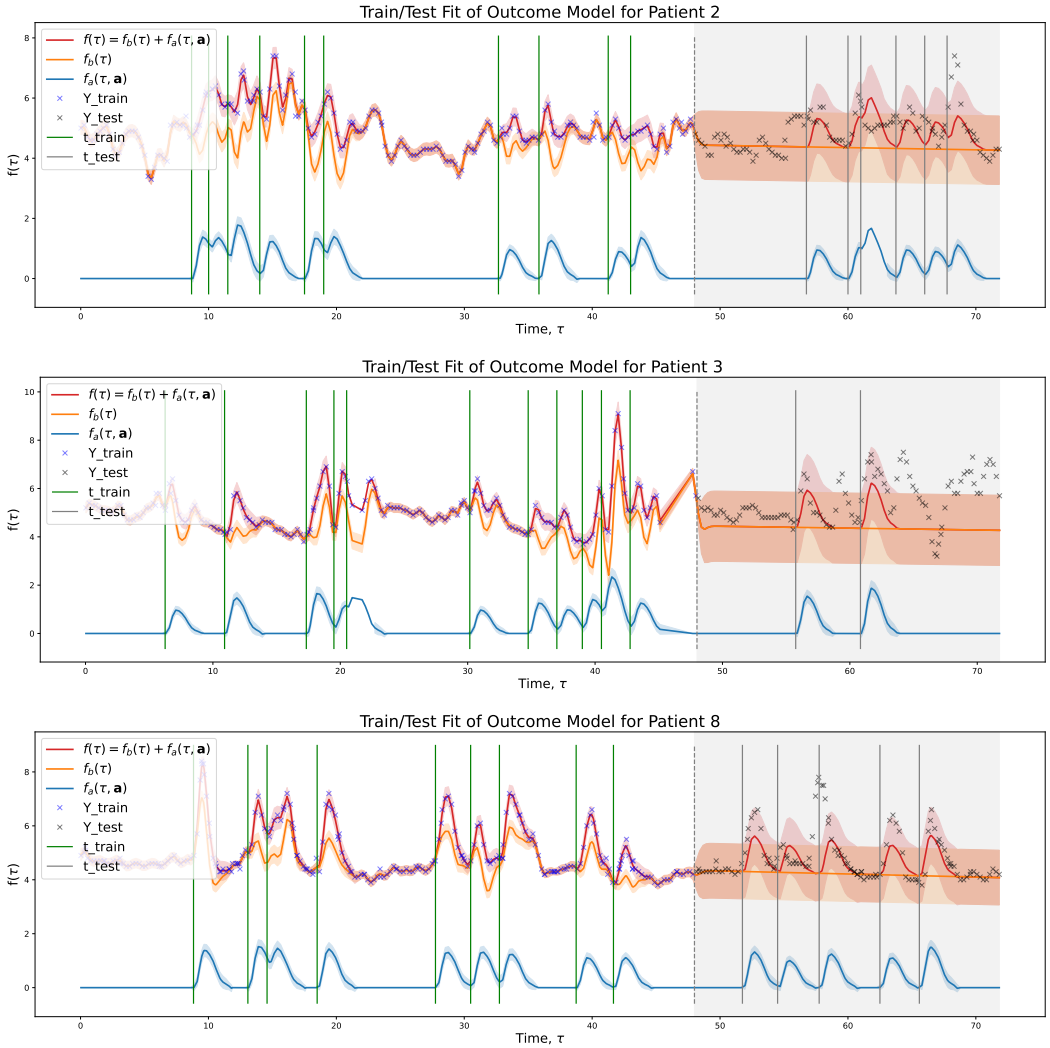

Figure 6: The learned ground-truth functions f_b and f_a for Patients {2,3,8}. The mean functions are the sums of (i) a linear mean close to a constant with a small decline, and (i) a short-term noise kernel [Ashrafi et al., 2021]. The treatment shape function is close to a Gaussian bump [Wyatt et al., 2021]. The treatment effect magnitudes are in the interval $[1.0, 2.5]$.

The learned time-dependent effect function $f_t^{(r)}(\tau; \mathbf{a})$, which represents the learned shape, is a smooth function similar to a Gaussian bump [Wyatt et al., 2021]. The individual-specific linear function $f_m^{(r)}(m_i)$, which hierarchically scale the magnitude of the response for the given treatment mark, provide a magnitude between a reasonable region $[1.0, 2.5]$.

Each baseline function $f_b^{(r)}(\cdot)$ is the sum of (i) a SE kernel with a low lengthscale value, which matches the findings in Ashrafi et al. [2021] about the constant blood glucose baseline and (ii) a linear function with a small decline, which matches our intuition for the simulation setup that the meal likelihood should increase over time if no meals take place, which is implicitly captured by the small decline in the baseline progression of three patients. For the simulations, we exclude the baseline SE kernel that captures the short-term noise, as our aim in this study is rather to evaluate our model on hypothetical scenarios, instead of capturing the short-term fluctuations in the blood glucose.

For the outcome noise variable, we use an independent Gaussian variable, $\epsilon \sim \mathcal{N}(0, 0.1^2)$. We sample the treatment mark variables that correspond to the meal sizes from a distinct uniform distribution for each patient group: $m_{gr_1} \sim \mathcal{U}(5.0, 8.5)$, $m_{gr_2} \sim \mathcal{U}(4.25, 7.5)$, $m_{gr_3} \sim \mathcal{U}(2.0, 3.5)$, so that the treatment response magnitudes across patient groups are within similar ranges.

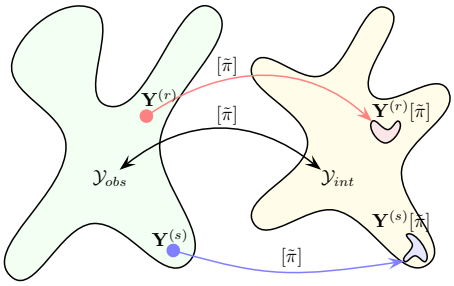

Figure 7: The intervention $[\tilde{\pi}]$ results in a stochastic mapping between the set of observational trajectories $\mathcal{Y}_{obs}$ and interventional trajectories $\mathcal{Y}_{int}$. A single discriminator is able to separate trajectories of $\mathcal{Y}_{obs}$ and $\mathcal{Y}_{int}$. Each counterfactual trajectory sample can be considered as a single interventional trajectory as well. Hence, a single discriminator can not separate interventional trajectories from counterfactual trajectories and we train a single, simple classifier for each counterfactual distribution $p(\mathbf{Y}^{(r)} \mid [\tilde{\pi}])$ of a single observed trajectory $\mathbf{Y}^{(r)}$.

We use the selected subsets of the treatment and outcome components detailed above to simulate samples from observational, interventional and counterfactual distributions. The observational data set contains 1-day treatment–outcome trajectories for 50 individuals. The interventional and counterfactual data sets are detailed in Appendix J.2.1, as they require the definition of the discriminator first. For each individual (r), we train estimated models on the observational data set to obtain estimated model components: $\mathcal{M}_{est}^{(r)} = \{\hat{\lambda}^{*}_{\pi^{(r)}}(\tau), \hat{f}_b^{(r)}(\tau), \hat{f}_a^{(r)}(\tau; \mathbf{a}), \hat{\epsilon}(\tau)\}$.

J.2.1 Metric: Discriminator Accuracy (DACC)

The marginal distribution of the outcome trajectory $\mathbf{Y}$ is not available in closed form (Equation 7), as it is not possible to integrate out treatments $\mathbf{a}$. However, our joint model allows for sampling observational, interventional and counterfactual outcome trajectories jointly with their treatments. A Monte Carlo integration can be performed to estimate the expected trajectories. However, in our experiments we found out that treatment effects are largely smoothed out by the Monte Carlo integration and hence the expected trajectory is not informative, since treatment effects in continuous time do not necessarily overlap in same time intervals for distinct trajectory samples.

To measure how similar predicted trajectories are to samples from a target ground-truth distribution, we train discriminators. Ideally, for samples of the same distribution, predicted trajectories should be inseparable from ground-truth trajectories, leading to a 50% discriminator accuracy (DACC). For example, the DACC should be close to 50% for predicted vs. ground-truth interventional samples (yellow region in Figure 7) or for predicted vs. ground-truth counterfactual samples (red region in Figure 7). Similarly, for samples from distinct target distributions, predicted trajectories should be separable from ground-truth trajectories, leading to a high DACC. For example, the DACC should be close to 100% for observational vs. interventional samples (green vs. yellow regions in Figure 7) or for interventional vs. counterfactual samples (yellow vs. red regions in Figure 7).

Classifying the ground-truth vs. predicted time-series is a non-trivial problem (or similarly, measuring the distance between them). There is no method that generalizes to all time-series data sets [Bagnall et al., 2017, Abanda et al., 2019, Ismail Fawaz et al., 2019]. Observational and interventional trajectories show a greater deal of variation in their dynamics compared to counterfactual trajectories, as they inherently have noise distributions with higher variance. As a result, observational and interventional trajectories contain temporary treatment effects initialized at different time points, even within samples from the same ground-truth distribution.

Shifts in treatment times make it impossible to measure the distance between two target time-series by a lock-step alignment method such as the mean squared error (MSE) [Abanda et al., 2019]. An elastic alignment method such as dynamic time warping (DTR) is a better distance candidate, as it is invariant to time shifts in meal times and provides a small distance value for shifted treatment responses. However, we would like our comparison metric to be as invariant as possible to time shifts due to the sampling noise, but as sensitive as possible to time shifts due to a distinct treatment

policy. For this reason, we employ a black-box discriminator function to separate interventional vs. interventional samples, parameterized by a neural network (NN).

For the neural net architecture, we experimented with three benchmark architectures provided in Ismail Fawaz et al. [2019]: a multilayer perceptron (MLP) classifier, a fully-connected 1-D convolutional classifier and a residual network classifier. We choose the MLP classifier as it provides higher DACC values for our setup. The architecture is straightforward with 3 hidden layers of 500 hidden units each. The neural network takes a pair of time-series samples as input. Each pair consists of a predicted and ground-truth trajectory. The neural network predicts a binary output to decide which one is fake and which one is real.

On the other hand, counterfactual trajectories show less variation and more distinctive features, due to the conditioning of noise variables to the observed data. To separate counterfactual vs. counterfactual samples, we use a simpler discriminator, which is a 1-nearest neighbour (1-NN) classifier using the MSE distance metric [Abanda et al., 2019].

Each classifier experiment is repeated 10 times with a different seed. For each seed, the DACC result is the average DACC value for 4 distinct policy-patient group combinations. For example, one combination can consist of a treatment policy A and an outcome group gr_2 with model components: $\{\pi, f_b, f_a\} = \{\pi_A, f_{b,gr_2}, f_{a,gr_2}\}$. For this group, an intervention on the policy label will set it to π_B: $[\tilde{\pi} = \pi_B]$. Below, we explain the train/validation/test sets for a single policy-patient group combination.

To create train/validation/test sets for the interventional distribution, we sample $400/100/100$ interventional outcome trajectories for a given patient, both with estimated and ground-truth model components. Outcome trajectory samples $\mathbf{Y}$ are 1-day long. We select the model with the highest accuracy on the validation set. On the test set, we test the ability of the discriminator to separate the remaining $100 + 100$ days of estimated vs. ground-truth time-series.

The counterfactual data set contains (i) 10 days of observational treatment-outcome trajectories and (ii) counterfactual samples that are conditioned on these days: 40/20 counterfactual outcome trajectories for train and test sets for each observational day. The resulting train/test set have the size of $10 \times \{40/20\}$. We train a single 1-NN discriminator for each observed day (10 discriminators in total) to separate estimated trajectories from ground-truth trajectories, as each observed day have a distinct counterfactual distribution. On the test set, we test the ability of each discriminator to separate $20 + 20$ counterfactual days of estimated vs. ground-truth time-series.

References

Odd O Aalen, Mats J Stensrud, Vanessa Didelez, Rhian Daniel, Kjetil Røysland, and Susanne Strohmaier. Time-dependent mediators in survival analysis: Modeling direct and indirect effects with the additive hazards model. *Biometrical Journal*, 62(3):532–549, 2020.

Amaia Abanda, Usue Mori, and Jose A Lozano. A review on distance based time series classification. *Data Mining and Knowledge Discovery*, 33(2):378–412, 2019.

Abubakar Abid, Mert Yuksekgonul, and James Zou. Meaningfully debugging model mistakes using conceptual counterfactual explanations. In *International Conference on Machine Learning*, pages 66–88. PMLR, 2022.

Ryan Prescott Adams, Iain Murray, and David JC MacKay. Tractable nonparametric bayesian inference in poisson processes with gaussian process intensities. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 9–16, 2009.

Mauricio Alvarez, David Luengo, and Neil D Lawrence. Latent force models. In *Artificial Intelligence and Statistics*, pages 9–16. PMLR, 2009.

Reza A Ashrafi, Aila J Ahola, Milla Rosengård-Bärlund, Tuure Saarinen, Sini Heinonen, Anne Juuti, Pekka Marttinen, and Kirsi H Pietiläinen. Computational modelling of self-reported dietary carbohydrate intake on glucose concentrations in patients undergoing roux-en-y gastric bypass versus one-anastomosis gastric bypass. *Annals of medicine*, 53(1):1885–1895, 2021.

Anthony Bagnall, Jason Lines, Aaron Bostrom, James Large, and Eamonn Keogh. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data mining and knowledge discovery*, 31(3):606–660, 2017.

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*, 2020.

Li-Fang Cheng, Bianca Dumitrascu, Michael Zhang, Corey Chivers, Michael Draugelis, Kai Li, and Barbara Engelhardt. Patient-specific effects of medication using latent force models with Gaussian processes. In *International Conference on Artificial Intelligence and Statistics*, pages 4045–4055. PMLR, 2020.

John Cunningham, Zoubin Ghahramani, and Carl Rasmussen. Gaussian processes for time-marked time-series data. In *Artificial intelligence and statistics*, pages 255–263. PMLR, 2012.

Vanessa Didelez. Causal reasoning for events in continuous time: A decision-theoretic approach. In *ACI @ UAI*, pages 40–45, 2015.

Vanessa Didelez, Philip Dawid, and Sara Geneletti. Direct and indirect effects of sequential treatments. *arXiv preprint arXiv:1206.6840*, 2012.

Tian Gao, Dharmashankar Subramanian, Debarun Bhattacharjya, Xiao Shou, Nicholas Mattei, and Kristin P Bennett. Causal inference for event pairs in multivariate point processes. *Advances in Neural Information Processing Systems*, 34:17311–17324, 2021.

Alan G Hawkes. Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2):193–198, 2018.

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

William Hua, Hongyuan Mei, Sarah Zohar, Magali Giral, and Yanxun Xu. Personalized dynamic treatment regimes in continuous time: A bayesian approach for optimizing clinical decisions with timing. *Bayesian Analysis*, 1(1):1–30, 2021.

Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. Deep learning for time series classification: a review. *Data mining and knowledge discovery*, 33(4):917–963, 2019.

ST John and James Hensman. Large-scale cox process inference using variational fourier features. In *International Conference on Machine Learning*, pages 2362–2370. PMLR, 2018.

PA W Lewis and Gerald S Shedler. Simulation of nonhomogeneous poisson processes by thinning. *Naval research logistics quarterly*, 26(3):403–413, 1979.

Siqi Liu and Milos Hauskrecht. Nonparametric regressive point processes based on conditional gaussian processes. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 1064–1074, 2019.

Chris Lloyd, Tom Gunter, Michael Osborne, and Stephen Roberts. Variational inference for gaussian process modulated poisson processes. In *International Conference on Machine Learning*, pages 1814–1822. PMLR, 2015.

Judith J Lok. Statistical modeling of causal effects in continuous time. *The Annals of Statistics*, 36 (3):1464–1507, 2008.

Alexander G de G Matthews, James Hensman, Richard Turner, and Zoubin Ghahramani. On sparse variational methods and the kullback-leibler divergence between stochastic processes. In *Artificial Intelligence and Statistics*, pages 231–239. PMLR, 2016.

Jesper Møller, Anne Randi Syversveen, and Rasmus Plenge Waagepetersen. Log gaussian cox processes. *Scandinavian journal of statistics*, 25(3):451–482, 1998.

Kimia Noorbakhsh and Manuel Gomez Rodriguez. Counterfactual temporal point processes. *arXiv preprint arXiv:2111.07603*, 2021.

Michael Oberst and David Sontag. Counterfactual off-policy evaluation with gumbel-max structural causal models. In *International Conference on Machine Learning*, pages 4881–4890. PMLR, 2019.

Yosihiko Ogata. On lewis' simulation method for point processes. *IEEE transactions on information theory*, 27(1):23–31, 1981.

Judea Pearl. *Causality*. Cambridge university press, 2009.

Judea Pearl and James M Robins. Probabilistic evaluation of sequential plans from causal models with hidden variables. In *UAI*, volume 95, pages 444–453. Citeseer, 1995.

Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer school on machine learning*, pages 63–71. Springer, 2003.

Jakob Gulddahl Rasmussen. Temporal point processes: the conditional intensity function. *Lecture Notes, Jan*, 2011.

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 Robins. Estimation of the time-dependent accelerated failure time model in the presence of confounding factors. *Biometrika*, 79(2):321–334, 1992.

James M Robins. Addendum to "a new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect". *Computers and Mathematics with Applications*, 14(9-12):923–945, 1987.

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

Pål Christie Ryalen, Mats Julius Stensrud, Sophie Fosså, and Kjetil Røysland. Causal inference in continuous time: an example on prostate cancer therapy. *Biostatistics*, 21(1):172–185, 2020.

Peter Schulam and Suchi Saria. A framework for individualizing predictions of disease trajectories by exploiting multi-resolution structure. *Advances in neural information processing systems*, 28, 2015.

Peter Schulam and Suchi Saria. Reliable decision support using counterfactual models. *Advances in neural information processing systems*, 30, 2017.

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.

Hossein Soleimani, Adarsh Subbaswamy, and Suchi Saria. Treatment-response models for counterfactual reasoning with continuous-time, continuous-valued interventions. *arXiv preprint arXiv:1704.02038*, 2017.

Stratis Tsirtsis and Manuel Gomez Rodriguez. Decisions, counterfactual explanations and strategic behavior. *Advances in Neural Information Processing Systems*, 33:16749–16760, 2020.

Stratis Tsirtsis, Abir De, and Manuel Rodriguez. Counterfactual explanations in sequential decision making under uncertainty. *Advances in Neural Information Processing Systems*, 34:30127–30139, 2021.

Patrick Wyatt, Sarah E Berry, Graham Finlayson, Ruairi O'Driscoll, George Hadjigeorgiou, David A Drew, Haya Al Khatib, Long H Nguyen, Inbar Linenberg, Andrew T Chan, et al. Postprandial glycaemic dips predict appetite and energy intake in healthy individuals. *Nature metabolism*, 3(4): 523–529, 2021.

Yanbo Xu, Yanxun Xu, and Suchi Saria. A bayesian nonparametric approach for estimating individualized treatment-response curves. In *Machine learning for healthcare conference*, pages 282–300. PMLR, 2016.

Guangyi Zhang, Reza A Ashrafi, Anne Juuti, Kirsi Pietiläinen, and Pekka Marttinen. Errors-in-variables modeling of personalized treatment-response trajectories. *IEEE Journal of Biomedical and Health Informatics*, 25(1):201–208, 2020.