Statistics in Medicine

WILEY

Statistics in Medicine

TUTORIAL IN BIOSTATISTICS OPEN ACCESS

Continuous-Time Causal Inference With Marked Point Process Weights: An Example on Sodium-Glucose Co-Transporters 2 Inhibitor Medications and Urinary Tract Infection

Sumeet Kalia[1,2] | Olli Saarela[3] | Tao Chen[2] | Braden O'Neill[2,3] | Christopher Meaney[2] | Rahim Moineddin[2,3] | Babak Aliarzadeh[2] | Frank Sullivan[4] | Michelle Greiver[2]

[1]Department of Statistics, University of Manitoba, Winnipeg, Canada | [2]Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada | [3]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada | [4]School of Medicine, University of St Andrews, St Andrews, UK

Correspondence: Sumeet Kalia (Sumeet.Kalia@umanitoba.ca)

Received: 20 April 2024 | **Revised:** 31 March 2025 | **Accepted:** 8 April 2025

Funding: This work was supported by the Natural Sciences and Engineering Research Council of Canada (Grant No. 534600).

Keywords: causal inference | constrained optimization | electronic health records | marginal structural models | marked-point process weights | primary care | time-to-recurrent event analysis

ABSTRACT

Treatment-confounder feedback is present in time-to-recurrent or longitudinal event analysis when time-dependent confounders are themselves influenced by previous treatments. Conventional models produce misleading statistical inference of causal effects due to conditioning on these factors on the causal pathway. Marginal structural models are often applied to quantify the causal treatment effect, estimated using longitudinal weights that mimic the randomization procedure by balancing the covariate distributions across the treatment groups. The weights are usually constructed in discrete time intervals, which is appropriate if the follow-up visits are scheduled and regular. However, in primary care, visit times can be irregular and informative, and the treatment history consists of duration and doses. This can be modeled through a continuous-time marked point process. We constructed a continuous-time marginal structural model to estimate the effect of cumulative exposure to Sodium-Glucose co-Transporters 2 Inhibitor (SGLT-2i) medications on time-to-recurrent urinary tract infection (UTI). We used a cohort of type II diabetes patients with chronic kidney disease and constructed a marked point process that characterized the recurrent flare episodes of primary care visits (i.e., point process) with marks for the multinominal dose (none, low, high) of SGLT-2i medications and recurrent episodes of UTI. We applied the stabilized and optimal treatment weights to estimate the hypothesized causal effect. Our results are concordant with earlier findings in which the recurrent episodes of UTI did not increase when patients were prescribed low dose or high dose of SGLT-2i medications.

Abbreviations: BMI, Body mass index; COPD, Chronic obstructive pulmonary disease; CKD, Chronic kidney disease; EHR, Electronic Health Records; FBG, Fasting blood glucose; HbA1c, Hemoglobin A1c; HDL, High Density Lipoprotein; ICD-9, International Classification of Diseases version-9; LDL, Low Density Lipoprotein; MCSD, Monte Carlo Standard Deviation; MSM, Marginal structural model; rMSE, root Mean Square Error; SGLT-2i, Sodium glucose co-transporter-2 inhibitors; UTOPIAN, University of Toronto Practice-Based Research Network; UTI, Urinary tract infection.

1 | Introduction

In longitudinal or recurrent event data analysis, time-dependent confounders (e.g., comorbidities or other treatments) exist as mediating variables that predict subsequent treatment and can themselves be influenced by previous treatment. This phenomenon is known as treatment-confounder feedback. The time-dependent treatment may, in turn, affect the outcome in two ways: as a direct effect of the treatment and/or as an indirect effect of the previous treatment mediated by a change in the time-dependent confounder. The estimation of marginal treatment effects is complicated by the presence of treatment-confounder feedback. For example, the adjustment of time-dependent confounders in the longitudinal causal structure may block some of the indirect pathways [1]. In contrast, nonadjustment of the time-dependent confounder may generate biased treatment estimates. This is partly attributed to *"confounding by indication"*, in which treatment is more likely to be assigned to subjects with a worse prognosis [2]. For this reason, conventional regression models may produce biased estimates of causal effects [3].

Instead of adjusting for time-dependent confounders, estimation of marginal structural models accounts for time-dependent covariates using inverse probability weights. In longitudinal causal inference, the inverse probability weights describe the estimated probability of the observed treatment history at time t, conditioned on observed history with respect to past treatment, confounders, and baseline covariates. Traditionally, the discretization of continuous time is employed to estimate the weights. As an example, Robins et al. [1] and Hernán et al. [4] describe weighted pooled logistic regression (assuming discrete time intervals) when quantifying the treatment effects. The approximation of the treatment effect in discretized time intervals using pooled logistic regression requires the probability of events to be rare within each time interval when the treatment effect is defined using the hazard ratio [5]. However, as demonstrated by Xiao et al. [3], the marginal structural Cox model may produce more accurate (i.e., reduced bias) and more precise (i.e., reduced standard error) estimates of causal effects when compared to pooled logistic regression approximation, especially when the time-to-recurrent outcome is frequent. Regardless of how the continuous-time scale might be discretized, Guerra et al. [6] demonstrated that the discretization of the longitudinal cohort not only modifies the definition for the causal parameter of interest but also results in loss of information for time-dependent confounders. In our context, the recurrent episodes of urinary tract infection (UTI) in type II diabetes population with history of chronic kidney disease (CKD) presenting irregularly to primary care clinics motivate the application of continuous-time causal framework [7].

In many population-based registries, medical information is updated on daily basis, and this necessitates the use of reduced granularity of the time intervals. The binning of event times in discrete time units is known to increase bias and variance of estimators [3]. While the marginal structural models using the discrete time intervals might be appropriate when follow-up visits in a longitudinal cohort are scheduled in advance, a more natural approach is to model the treatment assignment mechanism in continuous time [8]. If the individual measurement times fall in small discrete time intervals, the covariate dependent observation time may be reframed as a missing-data problem [9]. However, conceptualizing the potential outcome framework as a missing data problem becomes less trivial with continuous-time follow-up due to increase in the number of possible measurement times. Apart from this subtle difference, the continuous-time weights are conceptually similar to the inverse probability weights used in discrete-time intervals [10]. In continuous time, the covariate balance among treatment groups, which mimics the desirable features of randomization, can be described as a change of probability measure from an observational measure to hypothetical measure and these two measures are related through the Radon – Nikodym derivative which serves as a weight process [11].

The counting process for the occurrence of patient visits in primary care clinics may depend on the clinical evolution of patients, and this may introduce selection bias [12]. The selection bias may be characterized with the oversampling of events with poor health status [13]. Inverse-intensity weights are traditionally applied to correct for the informative visit process. However, this is not realistic due to the missing at random assumption in which we do not attempt to correct for the visit process to model a longitudinal binary outcome but rather combine the visit times and UTI status into a recurrent event outcome. Immediate treatment effect is difficult to estimate due to confounding by current health status, and thus we define the causal contrast using a delayed treatment effect under slightly weaker sequential randomization or stability assumption.

1.1 | Knowledge Gap

Less emphasis is placed in the statistical literature on the application of marginal structural models for time-to-recurrent outcomes in the presence of time-dependent confounding using continuous-time formulation [14]. Furthermore, to the best of our knowledge, continuous-time marked point processes have not been used to construct weights for the estimation of marginal structural models, although modeling the treatment process this way has been suggested by Hua et al. [15]. The purpose of optimized and stabilized weights is to emulate the necessary conditions of a randomized experiment (i.e., achieve appropriate balance for the distribution of time-dependent covariates across treatment groups). This article demonstrates the application of continuous-time marginal structural models with the constrained optimization of stabilized weights when assessing the cumulative treatment effect of sodium-glucose co-transporter 2 inhibitors (SGLT-2i) prescriptions on recurrent UTI events.

1.2 | Motivating Example

In some situations, randomized trials may not be feasible because they are expensive, untimely, or unethical [16]. Under these circumstances, large observational data repositories (e.g., collected from routine primary care clinics) provide an alternative source to emulate a target trial in which we can assess the long-term effects of medications (e.g., glucose lowering medications such as

SGLT-2i) on adverse clinical outcomes (e.g., UTI). The SGLT-2i medications are a new class of antidiabetic drugs with favorable clinical outcomes for blood glucose control, and reduced adverse outcomes related to cardiovascular and renal systems [17]. In contrast, the increased levels of glucosuria in urinary tract may increase the risk of genital infection, and increase the risk of UTI [18]. Although the increased presence of glucose in the urinary tract due to SGLT-2i medications has been consistently shown to increase the risk of genital infections [19], their associations with UTI is less clear, and earlier meta-analyses provide conflicting evidence [20]. To account for these concerns, regulatory bodies (e.g., the Therapeutic Goods Administration in Australia) requested the manufacturer of SGLT-2i medications to undertake a postmarket surveillance plan evaluating the risk of several adverse events including UTI. In this article, we carried out as-treated analyses for the prescribed treatment of SGLT-2i medications, and we define the causal contrast of interest as the change in the hazard of UTI among patients with cumulative exposure of low dose and high dose SGLT-2i medications. The time-dependent covariates are considered as the onset of various chronic comorbidities such as hypertension and depression.

2 | Methods

This section describes the framework for continuous-time marginal structural models. First, the causal identifiability assumptions are described, followed by the construction of a marked point process model to characterize the treatment process as a product of the visit intensity and the dose probabilities. The continuous-time stabilized weights and continuous-time optimal weights are formulated for marginal structural models for recurrent event intensity.

2.1 | Framework

A longitudinal cohort is considered for n individuals from time 0 to time $\mathcal{T}$ where individuals are indexed using $i = \{1, 2, \ldots, n\}$ and follow-up visits are indexed using j, and they take place in the continuous time. The history $\mathcal{F}_t$ encompasses all observed events, including baseline variables Z_i, covariate process $X_i(t)$, visit process $V_i(t)$, at-risk process $G_i(t)$, current dose process $D_i(t)$, cumulative dose process $A_i(t)$, and recurrent event outcome process $Y_i(t)$. We denote the intensity function of the recurrent event process as $\lambda_{Y_i}(t)$, covariate process as $\lambda_{X_i}(t)$, and visit process as $\lambda_{V_i}(t)$.

2.2 | Marked Point Process Weights

The marked point process is a generalization of the point process in which the random events of the point process are associated with a random vector (i.e., bivariate marks) [21]. For example, Mancini and Paganoni [22] used the marked point process to connect the temporal trend of hospitalization (the point process) with the length of stay (the mark) at each event. In this article, we used a bivariate marked point process to connect the visits to primary care clinic (the point process) with the multinomial dose of glucose-lowering medications and the occurrence of UTI events (bivariate marks).

We define a marked point process $(T_{ij}, \{D_{ij}, Y_{ij}\})$, where T_{ij} is the j^{th} observed visit time, D_{ij} is the SGLT-2i dose at that time, and Y_{ij} is the UTI status at that time. This marked point process deterministically implies the processes for multinomial current dose $D_i(t)$ and recurrent outcome $Y_i(t)$ conditioned on the occurrence of the visit $V_i(t)$. As an extension to earlier research [23], we conceptualized the marked point process in which the visit times T_{ij} are encoded as a point process and the dose of SGLT-2i prescription D_{ij} and recurrent UTI outcome Y_{ij} are encoded as the marks conditional on the visit times T_{ij}. The conditional dose probability (i.e., $P(D_i(s) = d_i(s)|\mathcal{F}_s)$) is of central importance as it represents the cumulative exposure of SGLT-2i medications (none, low, high) across longitudinal follow-up. We used the marked point process to specify the continuous-time weights for the treatment process as a product of visit weights and dose weights as

$$SW_i^A(t) = SW_i^V(t) \times SW_i^D(t) \tag{1}$$

where $SW_i^V(t)$ and $SW_i^D(t)$ are the stabilized weights for visit process and dose process (respectively); and collectively they characterize the continuous-time treatment weights $SW_i^A(t)$.

In similar spirit to Dawid et al. [24] and Røysland et al. [11], we express the superscript $\mathcal{O}$ and $\mathcal{E}$ for observational and experimental data generating mechanisms (respectively) for the causal quantities of interest. Using this framework, we express the observational and experimental visit intensity functions as

$$\lambda_{V_i}^{\mathcal{O}}(t)dt = P^{\mathcal{O}}(dV_i(t) = 1|\mathcal{F}_t) \tag{2}$$

$$\lambda_{V_i}^{\mathcal{E}}(t)dt = P^{\mathcal{E}}(dV_i(t) = 1|\mathcal{F}_t^*) \tag{3}$$

We note that the observational intensity function is conditioned on full history $\mathcal{F}_t$ while the experimental intensity function is conditioned on the partial history $\mathcal{F}_t^*$ with the exclusion of time-dependent covariates $X_i(t)$ [8]. Likewise, the observational and experimental intensity functions for the recurrent outcome of UTI episodes are characterized as

$$\lambda_{Y_i}^{\mathcal{O}}(t)dt = P^{\mathcal{O}}(dY_i(t) = 1|\mathcal{F}_{t-\vartheta}) \tag{4}$$

$$\lambda_{Y_i}^{\mathcal{E}}(t)dt = P^{\mathcal{E}}(dY_i(t) = 1|\mathcal{F}_{t-\vartheta}) \tag{5}$$

where both intensity functions are conditioned with respect to temporal latency ϑ denoted in history $\mathcal{F}_{t-\vartheta}$.

Using the $\mathcal{O}$ and $\mathcal{E}$ notation, we define the continuous-time stabilized visit weights as

$$SW_i^V(t) = \exp\left(\int_0^t (\lambda_{V_i}^{\mathcal{E}}(s) - \lambda_{V_i}^{\mathcal{O}}(s))ds\right) \times \prod_{s \le t} \left(\frac{\lambda_{V_i}^{\mathcal{E}}(s)}{\lambda_{V_i}^{\mathcal{O}}(s)}\right)^{dV_i(s)} \tag{6}$$

We note the functional form of stabilized weights in continuous time (Equation (6)) as the likelihood ratio between the marginal and the conditional density functions. The continuous-time likelihood ratio process in marginal structural models is based on the ratio of two martingale measures, namely the observational $\mathcal{O}$ measure characterized using the conditional intensity function $\lambda^{\mathcal{O}}(\cdot)$, and the experimental $\mathcal{E}$ measure (i.e., as though it was

a randomized trial) characterized using the marginal intensity function $\lambda^{\mathcal{E}}(\cdot)$. In a heuristic sense, the likelihood ratio can be expressed as a transformation from observational scenario into a hypothetical scenario [11]. The causal (or unconfounded) inference is then generated using the probability measure based on a hypothetical randomized trial, which is weighted appropriately using this continuous-time likelihood ratio process. Using the earlier conceptualization of marked processes, the dose weights can be presented as

$$SW_i^{D}(t) = \prod_{s \leq t} \left(\frac{P(D_i(s) = d_i(s) | \mathcal{F}_s^*)}{P(D_i(s) = d_i(s) | \mathcal{F}_s)} \right)^{dV_i(s)} \tag{7}$$

where $\mathcal{F}_t^*$ is the partial history with the exclusion of time-dependent covariates $X_i(t)$. We combine the stabilized visit weights and dose weights to express the continuous-time treatment weights (as shown in Equation (1)). The treatment weights $SW_i^A(t)$ emulate the randomization procedure as a change of probability measure from the observational $\mathcal{O}$ measure to the experimental $\mathcal{E}$ measure.

2.3 | Assumptions

In continuous time, the stability assumption implies that the future evolution of the confounder process (given the observed history) is independent of the current treatment regimen [24]. Using the stability assumption, we also rule out unmeasured confounders, and assume that the present treatment assignment does not depend on the future outcomes in the experimental $\mathcal{E}$ measure, conditioned on the history $\mathcal{F}_{t-\vartheta}$ [8, 24]. We weaken the stability assumption by assuming temporal latency ϑ in the outcome process as

$$\lambda_{X_i}^{\mathcal{O}}(t | \mathcal{F}_t) = \lambda_{X_i}^{\mathcal{E}}(t | \mathcal{F}_t) \tag{8}$$

$$\lambda_{Y_i}^{\mathcal{O}}(t | \mathcal{F}_{t-\vartheta}) = \lambda_{Y_i}^{\mathcal{E}}(t | \mathcal{F}_{t-\vartheta}) \tag{9}$$

where $\mathcal{F}_{t-\vartheta}$ is the full history up to and including time $t - \vartheta$. We also note that the randomized trial measure $\lambda_{Y_i}^{\mathcal{E}}(t | \mathcal{F}_{t-\vartheta})$ is connected to the observed trial measure $\lambda_{Y_i}^{\mathcal{O}}(t | \mathcal{F}_{t-\vartheta})$. The stability assumption is weakened with the inclusion of temporal latency ϑ to account for the delayed ascertainment of the UTI status due to the biological manifestation of the infection and also due to the irregularities in the primary care visits.

We assume sequential positivity using the observational dose probability $P^{\mathcal{O}}(D_i(s) = d_i(s) | \mathcal{F}_s)$ and experimental dose probability $P^{\mathcal{E}}(D_i(s) = d_i(s) | \mathcal{F}_s^*)$ as $P^{\mathcal{O}}(D_i(s) = d_i(s) | \mathcal{F}_s) > 0 \Rightarrow P^{\mathcal{E}}(D_i(s) = d_i(s) | \mathcal{F}_s^*) > 0$. In addition, we also assume sequential positivity for the visit intensity process as $\lambda_{V_i}^{\mathcal{O}}(t) > 0 \Rightarrow \lambda_{V_i}^{\mathcal{E}}(t) > 0$ [24]. The absolute continuous measure allows the empirical expectation of random variables to be simulated, and, heuristically, this means that the likelihood ratio can be represented as a transformation from the observational study into the hypothetical randomized trial [11]. It is assumed that no two counting processes can jump simultaneously in continuous time.

2.4 | Marginal Structural Model

The marginal intensity function $\psi_i(t; \theta)$ characterizes the instantaneous progression from one state to another at time t conditioned on current state [25]. We describe the marginal intensity function for UTI using the experimental data generating mechanism $\mathcal{E}$ as $\psi_i(t; \theta)dt = P^{\mathcal{E}}(dY_i(t) = 1 | \mathcal{F}_{t-\vartheta}^*; \theta)$, where θ parameterizes the covariate effect. In particular, we use a generalization of the counting process devised for time-to-event outcomes to characterize UTI as the recurrent events. We apply the Andersen–Gill extension of weighted Cox partial likelihood in which the occurrence of an event (e.g., UTI) does not prohibit a patient from visiting the clinic [26]. The recurrent events in Andersen–Gill model are conditionally independent (uncorrelated) given the covariates, and the recurrent events are parameterized based on the dependencies using some functional form of the previous history $\mathcal{F}_t$. The Andersen–Gill model is used as Markov process to estimate the overall treatment effect on the intensity of the occurrence of recurrent event when the dependencies are mediated through time-varying covariates [27].

We estimated the marginal structural model using marked point process weights with the aim of making the treatment independent of the covariates $X_i(t)$ after weighting. The time-dependent treatment weights are encoded as a product of cumulative-time visit weights and cumulative-time dose weights. We represent the weighted Cox partial likelihood as

$$L(\theta) = \prod_{i=1}^{n} \prod_{u \in [0,t)} \left(\frac{SW_i^A(u) \psi_i(u; \theta)}{\sum_{l=1}^{n} G_l(u) SW_l^A(u) \psi_l(u; \theta)} \right)^{SW_i^A(u) dY_i(u)} \tag{10}$$

where $G_l(\cdot)$ denotes the at-risk process and $SW_i^A(\cdot)$ represents the stabilized treatment weights. Unlike the Cox partial likelihood function, the weighted Cox partial likelihood function does not satisfy the information equality and thus requires the application of a robust sandwich estimator. We encourage the interested reader to refer to Aalen et al. [28] and Cook [21] for a detailed theoretical background on survival and event history analysis.

2.5 | Optimal Weights

Bias-variance trade-off is often encountered in statistical research where an increase in model complexity may reduce bias at the cost of inflating variance. The high variability of the estimated treatment effects using marginal structural models may be explained due to strong associations between the covariate process and the treatment process [3]. For example, strong associations between the covariate process and the treatment process may exacerbate the variance in which extreme weights dominate the weighted analysis in the pseudo-population. The stabilization, normalization, or truncation of extreme weights may reduce this variability at the expense of introducing bias. The optimal probability weights formulated as a constrained optimization problem may mitigate this trade-off whereby improvements can be made to the precision (reduced variance) without sacrificing the consistency of the estimator [29].

Marginal structural models require the sequential positivity assumption in which the treatment propensities are bounded away from zero and one. However, violation of sequential positivity assumptions are common in observational studies and this may generate erroneous inference. For example, Xiao et al. [3]

describe how the violation of the positivity assumption due to strong associations between the covariate process and the treatment process may increase the bias and the variance of the causal parameter estimate. Although the regression coefficients are consistent with the correct model specification, the marginal structural models may be inefficient when the weights are heavily skewed [30]. In the context of discrete-time intervals, Bembom and Van der Laan [31] truncated the inverse probability weight estimator to minimize the mean square error. In principle, similar arguments for truncating the continuous-time weights to minimize the mean square error may also be applied for continuous-time weights [32].

Santacatterina and Bottai [29] proposed optimal weights using a quadratic constrained optimization problem where the truncation with threshold ξ may reduce the variability due to extreme weights at the expense of introducing negligible bias. In particular, Santacatterina and Bottai [29] introduced a method to obtain optimal weights in which the Euclidean distance from target weights is minimized using the constraint that reduces the variance of the resulting weighted estimator. A similar approach was also described by Zubizarreta [33] where stable weights were estimated by solving a constrained optimization problem to minimize the variance of the weights under the constraint that the mean value of the covariates remains within a given tolerance. Henceforth, in a similar spirit to Santacatterina and Bottai [29], we introduce optimal probability weights in marginal structural models to improve the bias-variance trade-off using the pre-specified precision ξ. We define optimal weights as $\widehat{SW}_i^{A^*}(t) = \min(\delta, \widehat{SW}_i^{A}(t))$, where δ is the upper truncation limit (using prespecified precision ξ). In the simulation study, we specify the truncation limit δ using the simulated sample (i.e., time-invariant and not individual-specific). However, the truncation limit $\delta(t)$ was specified as time-variant (i.e., defined with respect to the risk set of UTI events) in the application section. We recommend the latter approach if reasonable computational resources can be allocated to the optimization of stabilized weights with respect to the unique configuration of the risk sets.

We describe the quadratic optimization problem as

$$\min_{\widehat{SW}_i^{A^*}(t)} \|\widehat{SW}_i^{A}(t) - \widehat{SW}_i^{A^*}(t)\|_2 \tag{11}$$

$$\text{subject to } \|\widehat{SW}_i^{A^*}(t) - \widehat{\overline{SW}}_i^{A}(t)\|_2^2 \leq \xi \tag{12}$$

$$\widehat{SW}_i^{A^*}(t) \geq 0 \tag{13}$$

where $\|\cdot\|_2$ denotes the Euclidean norm and $\widehat{\overline{SW}}_i^{A}(t)$ describes the mean stabilized weights at time t. We specify Equation (11) as the objective function in which the Euclidean distance is minimized between stabilized weights $\widehat{SW}_i^{A}(t)$ and optimal probability weights $\widehat{SW}_i^{A^*}(t)$. Equation (12) constrains the variance of the stabilized weights $\widehat{SW}_i^{A}(t)$ using the prespecified level ξ while Equation (13) constrains that the optimized weights $\widehat{SW}_i^{A^*}(t)$ to be non-negative. The optimal probability weights are based on constrained optimization to control the precision of weights to find the closest set of weights to the original (i.e., nontruncated) weights. In general, unbiased weighted estimators contain excessively low precision (or high variance). In a similar spirit to Santacatterina et al. [32], we specified the values of ξ parameter as prespecified level of precision. The smaller values of ξ correspond to increased precision and larger estimated values of the objective function (Equation (11)) while larger values of ξ induces greater bias of $\widehat{SW}_i^{A^*}(t)$ with respect to the target parameter.

3 | Simulation Study

3.1 | Data Generation

The times are drawn sequentially using exponential distribution for counting processes $V_i(t)$ and $X_i(t)$, updating the intensities according to the observed history of the processes measured so far. Once the visit times are simulated, multinomial dose $D_i(t)$ and binary outcome $Y_i(t)$ were drawn at visit times t. We specify a multinomial distribution to model the three dose assignment, indicating no treatment ($d = 0$), low dose treatment ($d = 1$) and high dose treatment ($d = 2$). We used the graph in Figure 1 to summarize the interdependencies between the stochastic processes, and how they evolved over time. Administrative censoring

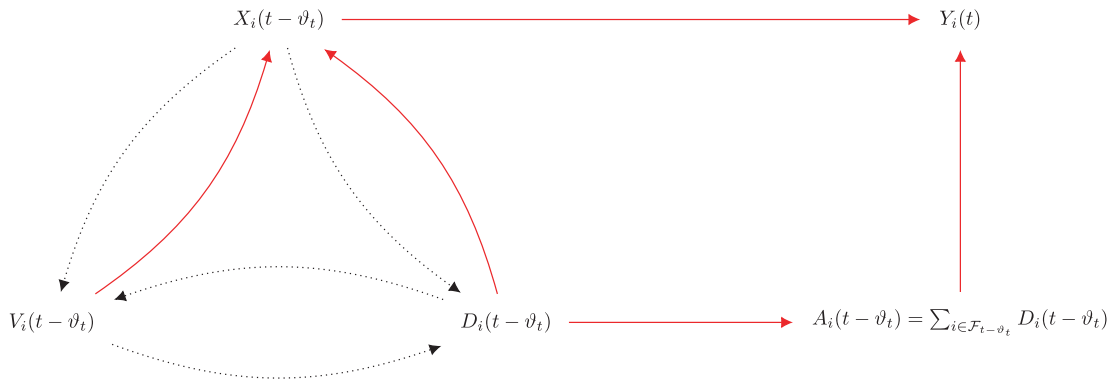

FIGURE 1 | Local independence graph for the observational $\mathcal{O}$ measure using the dotted edges and experimental $\mathcal{E}$ measure using the solid edges, where ϑ_t = temporal latency, $X_i(t-\vartheta_t)$ = covariate process, $V_i(t-\vartheta_t)$ = visit process, $D_i(t-\vartheta_t)$ = dose process, $A_i(t-\vartheta_t)$ = cumulative dose process, $Y_i(t)$ = outcome process, $\mathcal{F}_{t-\vartheta_t}$ = full history up to and including time $t-\vartheta_t$.

was imposed at $\tau = 6$, which denoted six years of study period, and random Type II censoring intensities were specified to be zero. The process for $X_i(t)$ was simulated to one event while the recurrent events for $V_i(t)$ was simulated with maximum number of 100 events for each patient within $\tau = 6$. We describe the total intensity of events as the sum of intensity functions as

$$\omega_i(t) = \lambda_{X_i}(t) + \lambda_{V_i}(t).$$

The following is the probability of the event type at time t, conditional that either event occurs at time t as

$$\left(\frac{\lambda_{X_i}(t)}{\omega_i(t)}, \frac{\lambda_{V_i}(t)}{\omega_i(t)}\right)$$

We specified the intensity function for the visit process as

$$\begin{aligned}\lambda_{V_i}(t) = \exp(&\eta_0 + \eta_1 X_i(t) \\ &+ \eta_2 \mathbb{1}(D_i(t) = 1) \\ &+ \eta_3 \mathbb{1}(D_i(t) = 2))\end{aligned}$$

where $\eta = \{\eta_0, \eta_1, \eta_2, \eta_3\}$. We define the intensity function for time-dependent covariate $X_i(t)$ as

$$\begin{aligned}\lambda_{X_i}(t) = \exp(&\zeta_0 + \zeta_1 \mathbb{1}(D_i(t) = 1) \\ &+ \zeta_2 \mathbb{1}(D_i(t) = 2))\end{aligned}$$

where $\zeta = \{\zeta_0, \zeta_1, \zeta_2\}$. By conditioning on the visit process, we draw the realization for multinomial dose ($d_i(t) \in \{0, 1, 2\}$) as

$$\begin{aligned}&P(D_i(t) = d_i(t) | dV_i(t) = 1, X_i(t); \phi) \\ &= \begin{cases} \frac{1}{1+\sum_{k\in\{1,2\}} \exp(\phi_{0k} + \phi_{1k} X_i(t))}, & d_i(t) = 0 \\ \frac{\exp(\phi_{0d} + \phi_{1d} X_i(t))}{1+\sum_{k\in\{1,2\}} \exp(\phi_{0k} + \phi_{1k} X_i(t))}, & d_i(t) \neq 0 \end{cases}\end{aligned} \tag{14}$$

where $\phi = \{\phi_{0k}, \phi_{1k}, \phi_{0d}, \phi_{1d}\}$. Together the dose assignment $D_i(t)$ and visit process $V_i(t)$ characterize the cumulative treatment process $A_i(t)$. This parameterization of the marked point process incorporates treatment-confounder endogeneity, in which the past treatment influences the current covariate, which in turn modifies the future treatment [34]. In this context, the treatment-confounder feedback (or endogeneity) is mediated through the visit process and the dose process to characterize the treatment process. Furthermore, by conditioning on the visit process, we draw the realization for binary outcome as

$$\begin{aligned}&P(Y_i(t) = 1 | dV_i(t) = 1, A_{i1}(t - \vartheta_t), A_{i2}(t - \vartheta_t), X_i(t)) \\ &= \text{expit}(\theta_0 + \theta_1 A_{i1}(t - \vartheta_t) \\ &\quad + \theta_2 A_{i2}(t - \vartheta_t) \\ &\quad + \theta_3 X_i(t))\end{aligned} \tag{15}$$

where $\Theta = \{\theta_0, \theta_1, \theta_2, \theta_3\}$ encodes the conditional probability of recurrent outcome. The transient duration of low dose and high dose prescription is determined as the time-interval between two consecutive visit times when the prescription might be discontinued at the latter visit, and the time-intervals for active prescriptions are then accumulated with respect to study period to define cumulative dose. In Equation (15), $A_{i1}(t - \vartheta_t)$ defines the cumulative low dose and $A_{i2}(t - \vartheta_t)$ defines the cumulative high dose, where the temporal latency ϑ_t is indexed with respect to time t.

3.2 | Marginal Effect

Earlier we characterized the intensity functions using the conditional data generating mechanism $\lambda^{\mathcal{O}}(\cdot)$, and now we construct the marginal intensity functions $\lambda^{\mathcal{E}}(\cdot)$ with respect to $X_i(t)$ to calculate the true value of the marginal effect. The marginal intensity function might differ from the conditional intensity function in which the proportionality assumption may be satisfied for the conditional intensity but not for the marginal intensity [23]. If we consider "no-confounding" scenario in which the covariate process does not modify the time-dependent treatment process, then we can characterize the randomized trial measure $\mathcal{E}$ using observational trial measure $\mathcal{O}$. Since the treatment effect is mediated through the visit intensity and dose response, the no-confounding scenario requires the coefficients for $X_i(t)$ to be set to zero in the visit intensity and the conditional probability of the dose assignment (see Table 1). Figure 1 describes the no-confounding scenario except with the absence of the dotted edges into the visit process and the dose process. The no-confounding scenario using $\mathcal{E}$ measure is then used to calculate the true value of the cumulative marginal effect of the low dose and high dose exposure. This marginal effect is used as the parameter value when estimating the treatment effect under the $\mathcal{O}$ measure (i.e., confounding scenario).

3.3 | Estimation of Treatment Weights

The visit weights $\widehat{SW}_i^V(t)$ are estimated using the Andersen-Gill extension of Cox model:

$$\widehat{SW}_i^V(t) = \frac{\exp\left\{-\int_0^t \lambda_{V_i}^{\mathcal{E}}(u;\ \widehat{\eta}^*)du\right\}}{\exp\left\{-\int_0^t \lambda_{V_i}^{\mathcal{O}}(u;\ \widehat{\eta})du\right\}} \times \prod_{s\le t}\left\{\frac{\lambda_{V_i}^{\mathcal{E}}(s;\ \widehat{\eta}^*)}{\lambda_{V_i}^{\mathcal{O}}(s;\ \widehat{\eta})}\right\}^{dV_i(s)} \tag{16}$$

A product (i.e., $\prod_{s\le t}$) of marginal and conditional intensities is taken after the first visit at time t. The dose weights are estimated using the multinomial model $\widehat{SW}_i^D(t)$ with the parameterized basis spline function in marginal and conditional probabilities as

$$\widehat{SW}_i^D(t) = \prod_{s\le t}\left(\frac{P(D_i(s) = d_i(s);\ \hat{\phi}_d^*)}{P(D_i(s) = d_i(s) | X_i(s); \hat{\phi}_d)}\right)^{dV_i(s)} \tag{17}$$

The stabilized treatment weights were combined as a product of visit weights and dose weights with respect to the unique configuration of the risk set as $\widehat{SW}_i^A(t) = \widehat{SW}_i^V(t) \times \widehat{SW}_i^D(t)$. In Equation (16) and (17), the product notation is used to represent the unique configuration of the risk sets arranged in chronological order. For example, consider a right-censored study with three individuals and a six-year follow-up period, in which two individuals experienced an event at 1.7 and 2.5 years of follow-up. In this example, the product is defined with the following configuration of the risk set arranged chronologically: (0, 1.7], (1.7, 2.5], (2.5, 6].

TABLE 1 | Parameter specification of the simulation study for the observational and experimental data generating mechanism.

Notation	Description	Parameter value
Data generating mechanism		
$\lambda_{V_i}(t)$	Visit intensity	$\eta_0 = -3, \eta_1 = 0.5; 1.0; (0^*)$
		$\eta_2 = 0.5; (0^*), \eta_3 = 0.5; (0^*)$
$P(D_i(t)\|dV_i(t) = 1, [\cdot])$	Dose probability	$\phi_{01} = 0, \phi_{02} = 0$
		$\phi_{11} = 0.5; 1.0; (0^*), \phi_{12} = 0.5; 1.0; (0^*)$
$\lambda_{X_i}(t)$	Covariate process	$\zeta_0 = -1, \zeta_1 = 0.1$
		$\zeta_2 = 0.1$
$P(Y_i(t)\|dV_i(t) = 1, [\cdot])$	Outcome probability	$\theta_0 = 0, \theta_1 = -0.25$
		$\theta_2 = -0.50, \theta_3 = 1.00$
Parameters of simulation study		
ξ	Precision of stabilized weights	0.001
τ	Maximum study follow-up time (years)	6
N	Number of subjects simulated	200

Note: * indicates no confounding scenario (experimental data generating mechanism). [·] indicates parameterized with respect to the covariate process and other processes (refer back to main text).

Initially, the inverse probability weights are calculated separately for each risk set, and then a cumulative product is taken in chronological order.

Constrained optimization was then performed on the stabilized treatment weights to estimate the optimal weights under the prespecified precision ξ. Since the objective function (11) was convex and constraints were linear, the numerical integration produced a unique solution. The optimal weights were used to produce estimators with the minimum bias among a set of estimators with the same specified precision ξ using a constrained nonlinear optimization problem (e.g., Euclidean norm). A quantile function ranging from zero to one in increments of 0.1 was specified for stabilized treatment weights in the objective function in Equation (11) to reduce the computational burden. We used the `CVXR` package in R with splitting conic solver (SCS) to optimize the stabilized treatment weights [35].

3.4 | Estimation Procedure

After the estimation of stabilized weights and their optimal counterparts using the prespecified ξ, we fit the marginal structural models using the Andersen–Gill regression with the continuous-time treatment weights. Each individual may contribute a cluster of repeated person moments in the simulated data file. We used the sandwich estimator for Andersen–Gill regression to account for this clustering of repeated observations. Following the earlier guidelines on Monte Carlo simulations in medical research [36], we specify the combinations of the simulation parameters in Table 1. The Monte Carlo simulation study assessed the finite sample properties using the inferential metrics of mean bias, Monte Carlo Standard Deviation (MCSD), root mean square error (rMSE), and coverage probabilities [36]. We replicated the data generation with experimental $\mathcal{E}$ measure and observational $\mathcal{O}$ measure with 1,000 iterations and with the sample size of 200 patients.

3.5 | Results

The treatment weights for $\mathcal{E}$ measure were concentrated near one with the following summary statistics: minimum=0.000; maximum=35.000; mean=1.085; median=1.052. Increased spread in treatment weights from one was observed in the presence of confounding. For example, the distribution of the optimized treatment weights in scenario one had the following characteristics: minimum=0.000; maximum=166.000; mean=1.353; median=1.062, while the distribution of the optimized treatment weights had the following characteristics for scenario two: minimum=0.000; maximum=12200.000; mean=26.606; median=1.3455.

Table 2 describes the results of the simulation study using unweighted sample, stabilized weighted sample, and optimized weighted sample. The parameter value for the marginal treatment effect is derived using the $\mathcal{E}$ measure in which the time-dependent treatment assignment mechanism (as a product of visit process and dose process) is exogenous. The performance metrics of bias, MCSD, rMSE, and coverage are then evaluated using the marginal treatment effect of low dose and high dose. All replicates (out of 1000 iterations) successfully converged for scenario 1, while 2 replicates failed to converge for scenario 2 (due to positivity violations with the increased magnitude of confounding). In comparison to the weighted samples, the unweighted samples had the largest bias, increased rMSE, and poor coverage probabilities. The stabilized weights had lower MCSD when compared to the optimized weights. Specifying the truncation limit to estimate the optimized weights sustained the similar magnitude of bias, rMSE, and coverage when compared with the stabilized weights across the two scenarios.

TABLE 2 | Performance metrics of the simulation study with nominal level of 0.95 for confidence interval coverage.

Type	Estimator	Parameter	Mean	Rel. bias (%)	MCSD	rMSE	Coverage[b]
Scenario 1: Visit and dose coefficients for $X_i(t)$ set to 0.5							
Unweighted	Low Dose	−0.293	0.040	−113.6	0.246	0.414	0.452
Unweighted	High Dose	−0.464	−0.108	−76.7	0.305	0.469	0.477
Stab Weights	Low Dose	−0.293	−0.335	14.4	0.288	0.291	0.883
Stab Weights	High Dose	−0.464	−0.493	6.2	0.338	0.339	0.841
Optimized weights	Low Dose	−0.293	−0.321	9.6	0.288	0.289	0.875
Optimized weights	High Dose	−0.464	−0.480	3.3	0.339	0.339	0.840
Scenario 2[a]: Visit and dose coefficients for $X_i(t)$ set to 1.0							
Unweighted	Low Dose	−0.293	0.096	−132.9	0.148	0.417	0.211
Unweighted	High Dose	−0.464	−0.036	−92.2	0.184	0.466	0.244
Stab Weights	Low Dose	−0.293	−0.291	−0.7	0.200	0.200	0.903
Stab Weights	High Dose	−0.464	−0.450	−3.0	0.245	0.245	0.864
Optimized weights	Low Dose	−0.293	−0.271	−7.5	0.200	0.202	0.884
Optimized weights	High Dose	−0.464	−0.432	−7.1	0.243	0.245	0.854

Abbreviations: MCSD = Monte Carlo Standard Deviation; Mean = mean estimate; Rel. Bias = Relative bias (in %); rMSE = root Mean Square Error.

[a]Two replicates failed to converge in Scenario 2.

[b]Coverage is estimated only using the successfully converging replicates.

4 | Application

At first, we describe the generation of a longitudinal cohort of type II diabetes patients with recurrent episodes of UTI. This is followed with the description of causal analyses to estimate the cumulative exposure of SGLT-2i medications on the incidence of UTI in the patient population with type II diabetes and CKD. We fit the Andersen–Gill extension of the Cox model for primary care visits, and multinomial model for the dose–response (none, low, high) of SGLT-2i medications in the female and male cohort. The treatment weights are then estimated and optimized as a product of visit weights and dose weights. The outcome model for UTI is used to assess the cumulative effect of low dose or high dose of SGLT-2i medication with 90 days of latency.

4.1 | Cohort Generation

We constructed a retrospective open cohort from January 01, 2015 to December 31, 2020 for patients in the primary care repository of the University of Toronto Practice-Based Research Network (UTOPIAN). The cohort membership with respect to study period was defined using the eligibility criteria and the censoring mechanism. Patients were enrolled in the open cohort when the following eligibility conditions were satisfied: (i) the patient was at least 18 years of age, (ii) the patient had at least one visit recorded in the billing table between January 1st 2015 and December 31st 2020, (iii) the patient satisfied the research quality criteria [37], (iv) the patient had an indication for type II diabetes and CKD in primary care EHRs [38]. All four conditions of the inclusion criteria must be met for the patient to be enrolled in the longitudinal cohort. We restricted the analyses to this population because SGLT-2i medications were appropriate for type II diabetes patients with history of CKD [39]. We imposed a censoring mechanism when at least one of the following criteria was satisfied: (i) patient was censored on June 30 when deceased year was recorded [40], (ii) patient was censored using the latest billing record when it fell outside the last two years of study period [41], (iii) if the first two censoring conditions were not satisfied then the patient was administratively censored at the end of the study period (i.e., December 31, 2020).

UTI events were identified from primary care EHRs using the diagnostic International Classification of Diseases version-9 (ICD-9) code {`590, 595, 599`} in the billing table [42]. Since the family physician may have prescribed the antimicrobial agents for UTI for up to ten days depending on the severity of the infection [7], we defined a washout period of three weeks to distinguish between unique (recurrent) UTI events over the study follow-up. Any UTI events prior to the eligibility criteria or after the censoring condition were not included in the retrospective open cohort. We stratified the analyses by biological sex because we hypothesized that the recurrent outcome of UTI was likely to be different due to human physiology [43]. We identified the SGLT-2i medications using the Anatomical Therapeutic Chemical (ATC) codes as detailed in the earlier work by Greiver et al. [44]. One of the effects of these medications was reduced re-absorption of glucose in the renal tubules (and increased excretion of glucose as a result), known as glucosuria; as a result, SGLT-2i medications may have increased the risk of genital infection and increased the risk of UTIs [18, 45]. Using the earlier meta-analysis [46], we categorized the SGLT-2i medication into high dose and low dose based on the following active ingredients: (i) Canagliflozin: {50mg, 100mg, 150mg} vs {300mg}; (ii) Dapagliflozin: {5mg} vs {10mg}; (iii) Empagliflozin: {5mg, 10mg, 12.5mg} vs {25mg}. We defined the transition probability to evolve in time for three states (no dose, low dose, high dose)

for SGLT-2i medications. The discontinuation of SGLT-2i medications was defined using a combination of the information available in the medication table (in order of precedence): (i) stop date, (ii) total refills, and (iii) duration count. In the presence of complete information, medication length was determined as a product of `RefillCount`, `DurationCount` (standardized as "days") and `DurationUnit` as further detailed in the data dictionary elsewhere [47] and in earlier work [48–50]. Depending on the available information in the medication table, we defined each SGLT-2i prescription to have minimum exposure of 30 days and maximum exposure of 365 days. We defined the cumulative duration of low dose or high dose SGLT-2i medication with 90 days of latency as an exposure when assessing the causal effect.

Age, biological sex, and neighborhood-level income quintiles were specified as baseline characteristics. We defined several comorbidities as time-dependent covariates: (i) chronic obstructive pulmonary disease (COPD), (ii) dementia, (iii) depression, (iv) dyslipidemia, (v) epilepsy, (vi) herpes zoster, (vii) hypertension, (viii) osteoarthritis, with the operational definitions described elsewhere [51]. Additional patient-specific characteristics included body mass index (BMI) and laboratory information on hemoglobin A1c, fasting blood glucose, high density lipoprotein (HDL), low density lipoprotein (LDL), triglycerides, and hemoglobin. In the causal diagram, the time-dependent comorbidities were specified as a common cause for the SGLT-2i treatment and recurrent events of UTI.

4.2 | Local Independence Graph

In a similar spirit to that of Ryalen et al. [10], we constructed a local independence graph (see Figure 2) for recurrent outcomes of UTIs. Local independence graphs are a class of causal graphs in which the intensities of certain events are independent of some but not all events in the past [52]. This graphical representation was used to describe the hypothetical randomization strategy in two step process: (i) treatment was assigned at the baseline and (ii) treatment initiated with new prescription thereafter. The dotted lines in Figure 2 described the hypothesized relationship between variables in the primary care data repository, but these dependencies (especially between SGLT-2i prescription and treatment initiation) would not be expected in a hypothetical randomized trial setting. The censoring node in Figure 2 was connected to all other transient states but complete arrows were not drawn for aesthetic reasons. The censoring mechanism was assumed to be locally independent of each and every node in Figure 2 while conditioning on the remaining nodes (i.e., blocking all back-door paths). With exception to biological sex and geographical characteristics (i.e., rurality and income quintiles), we noted that the covariate set in the vertices containing the baseline covariates (e.g., age, body mass index, diagnostic conditions, laboratory characteristics) evolved with time. We also noted that there were vertices from recurrent UTI events into future prescription for SGLT-2i medications as $\Delta t \to 0$. We imagined the local independence graph in Figure 2 to be continuously evolving over time, in which the edges may appear or disappear as more clinical information was collected in the primary care repository.

The purpose of local independence graph shown was to elucidate how variables influenced one another with respect to time. Since the cause must precede the effect, we described the causal effect between SGLT-2i prescription and UTI as an evolution in the stochastic processes rather than correlation between fixed variables [28]. The idea of counterfactual time-to-recurrent models was to capture a change in counting process with the progression

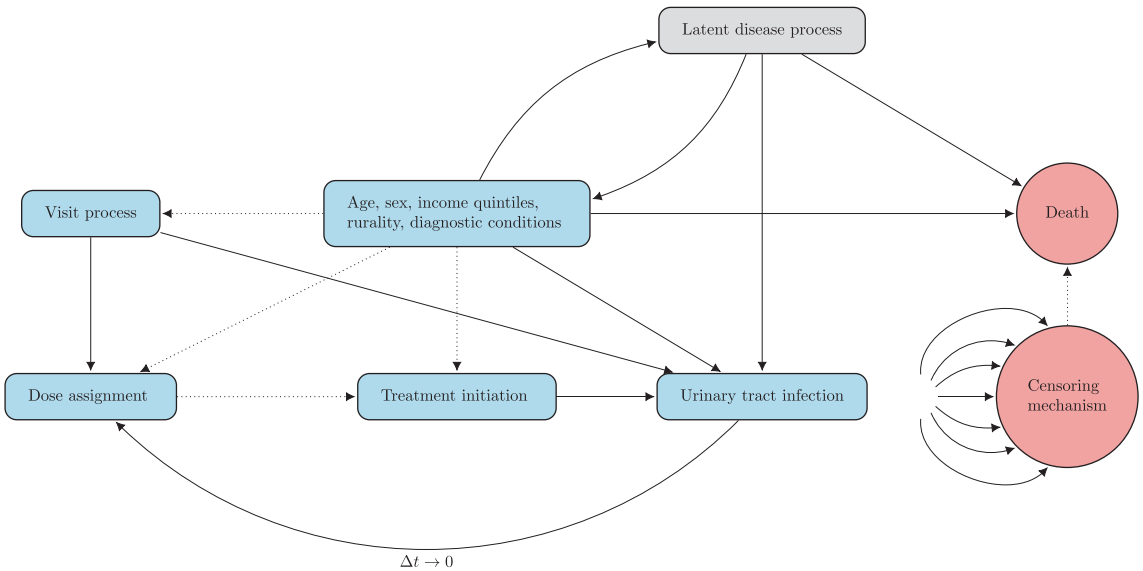

FIGURE 2 | Local independence graph for sodium glucose co-transporter-2 inhibitors (SGLT-2i) prescriptions and urinary tract infection (UTI). The absorbing states are denoted using the censoring mechanism and death (as red circle), while latent disease process is denoted using the gray rectangle. All transient states are described using the blue rectangles. Solid arrows describe temporal dependencies while dotted lines denote the latent dependencies in the primary care repository (not expected in hypothetical randomized trial). The back arrow from UTI to SGLT-2i prescription describes how the two processes may influence each other in the future (with $\Delta t \to 0$). Complete arrows from all transient states (for the blue rectangles) are not drawn toward censoring mechanism for aesthetic reasons.

of time defined over arbitrary risk set. We described the latent disease process in Figure 2 as a random effect in time-to-recurrent analyses. The local independence graph in Figure 2 provided a simplified mechanistic model for the treatment effect of SGLT-2i medications, and how it related to the onset of UTI. It was obvious that this causal representation in Figure 2 was somewhat superficial in a biological sense because it did not encapsulate the appropriate signaling pathways at the molecular level. Of course, a much more detailed understanding can be derived if we could specify biological pathways that might be responsible for increased onset of UTI. Nevertheless, it was necessary to acknowledge that the depth of the causal analysis applied in Figure 2 was not any greater than that permitted by the primary care EHRs. Therefore, depending on the available information, the resulting causal model was highly provisional, and it was subject to various confounders for a given mechanistic model. This conceptualization was dependent on the level at which we were operating, and in light of additional knowledge, the causal pathways may have dissolved into a number of other plausible relationships with more elaborative measurements.

4.3 | Marginal Structural Models

We applied the framework of marginal structural models to account for time-dependent covariates $X_i(t)$ between outcome process $Y_i(t)$ and treatment process $A_i(t)$. The marginal structural models used the stabilized and optimal continuous-time weights to compare the hazard of UTI as time-to-recurrent outcome. In this observational cohort, we needed to control for the following biases: (i) assigned treatment varied across subgroup of population and (ii) the dose–response at which the individual started the treatment. We applied the continuous-time optimal weights to mimic randomization for a scenario in which both the treatment using SGLT-2i prescriptions and time-to-treatment initiation were randomized. The marginal structural models compared the UTI episodes under the counterfactual scenario where a particular treatment (e.g., cumulative low dose SGLT-2i) was imposed to the entire population as oppose to its counterpart (e.g., cumulative high dose SGLT-2i). Using the Andersen–Gill model, we assumed the UTI events were conditionally independent (i.e., uncorrelated) for the given covariates with the following parameterization:

$$\begin{aligned}\lambda_Y(t) = \lambda_0(t)exp(&\theta_1 \times \text{age group}_i \\ &+ \theta_2 \times \text{income quintile}_i \\ &+ \theta_3 \times \text{cum}_{t-\vartheta}(\text{low dose SGLT-2i prescription}_i) \\ &+ \theta_4 \times \text{cum}_{t-\vartheta}(\text{high dose SGLT-2i prescription}_i)) \quad (18)\end{aligned}$$

where $\lambda_0(t)$ denoted the baseline hazard, and cum$(\cdot)$ denoted the cumulative duration of treatment (in years) until time $t-\vartheta$. In Equation (18), age group and income quintiles are the baseline characteristics. The temporal latency (ϑ) for the cumulative duration of the treatment was specified to be 90 days with respect to a visit at time t. Using the framework of marginal structural models, the parameters of interest are $\{\theta_3, \theta_4\}$ (in Equation 18), expressing the cumulative effect of low-dose and high-dose SGLT-2i medications with 90 days of temporal latency for the recurrent outcome of UTI events, respectively. Robust standard error was used to account for subject-specific correlation in the person moments belonging to the same patient.

4.3.1 | Treatment Weights

We constructed the treatment weights $SW_i^A(t)$ using marked-point process in two components. First, we applied the Andersen–Gill model to estimate the marginal and conditional visit intensities. Second, after conditioning on the presence of primary care visits, we used the multinomial regression to estimate the marginal probability $\hat{p}_{is}^{*d}$ and conditional probability $\hat{p}_{is}^{d}$ for the dose–response of SGLT-2i prescriptions. The marginal probabilities were defined with respect to the exclusion of time-dependent covariates $X_i(t)$ (as listed earlier). As a whole, we expressed the treatment weights as the product of cumulative visit weights $SW_i^V(t)$ and cumulative dose weights $SW_i^D(t)$ as

$$\begin{aligned}SW_i^{\bar{A}}(t) &= \prod_{s\le t} SW_i^V(s) \times SW_i^D(s) \\ &= \left\{ \prod_{s\le t} exp\left(\int_0^t \lambda_{V_i}^{\mathcal{E}}(s) - \lambda_{V_i}^{\mathcal{O}}(s) ds \right) \right. \\ &\quad \left. \times \left[\frac{\lambda_{V_i}^{\mathcal{E}}(s)}{\lambda_{V_i}^{\mathcal{O}}(s)} \right]^{dN_V(s)} \right\} \times \left(\prod_{s\le t} \frac{\hat{p}_{is}^{*d}}{\hat{p}_{is}^{d}} \right)^{\mathbb{1}(V_{i(s)}=1)}\end{aligned}$$

In the absence of subsequent visit, the longitudinal information on treatment was only encoded using the baseline treatment. In such cases, the dose weights were one in the absence of visit. The continuous-time treatment weights $SW_i^{\bar{A}}(t)$ accounted for confounding bias due to the choice of treatment, and also accounted for the time-to-treatment initiation. After constructing the pseudo-population with respect to the time-dependent covariates, we estimated the effect of SGLT-2i dose on the recurrent events of UTI. We estimated the optimal weights (using the unique configuration of the risk sets) to handle extreme weights arising due to cumulative-time product weights. We applied the quadratic optimization problem listed in Section (2.5), and the detailed description was provided in the Appendix Section.

4.4 | Results

The complete set of descriptive results for the prevalence and incidence of UTI, along with visit model, treatment model, and outcome model, can be found in the Appendix Section. Herein, we highlighted the most relevant results for the cumulative effect of SGLT-2i medications on recurrent episodes of UTI using the optimal weights (see Table 3). In both male and female cohorts, patients with hypertension, osteoarthritis, COPD, and depression had increased hazard for primary care visit. Male patients in older age groups had increased hazard for primary care visits than younger age groups. In both male and female cohorts, we observed a decrease in hazard ratio for UTI with low or high cumulative dose of SGLT-2i medication. These findings supported the results from earlier meta-analysis in which no dose dependence of SGLT-2i medication was observed for several clinical adverse event including UTI [46].

TABLE 3 | Hazard ratios for urinary tract infection in male and female cohort using Andersen-Gill model.

Contrast	Hazard ratio	95% CI[b]		Z-score	P-value
Optimized weight male cohort					
Age group: 50-59 years vs 18-49 years	1.745	0.790	3.856	1.377	0.168
Age group: 60-69 years vs 18-49 years	0.928	0.454	1.898	−0.204	0.838
Age group: 70-79 years vs 18-49 years	1.193	0.593	2.397	0.495	0.621
Age group: 80+ years vs 18-49 years	1.517	0.760	3.029	1.181	0.237
Income Quintiles: 2 vs 1 (lowest)	1.082	0.853	1.373	0.649	0.516
Income Quintiles: 3 vs 1	0.859	0.662	1.115	−1.143	0.253
Income Quintiles: 4 vs 1	0.989	0.759	1.289	−0.082	0.935
Income Quintiles: 5 (highest) vs 1	0.885	0.684	1.145	−0.931	0.352
Cumulative low dose SGLT-2i medication[a]	1.011	0.611	1.672	0.041	0.967
Cumulative high dose SGLT-2i medication[a]	0.330	0.156	0.700	−2.891	0.004
Optimized weight female cohort					
Age group: 40-49 years vs 18-39 years	0.445	0.150	1.317	−1.463	0.144
Age group: 50-59 years vs 18-39 years	0.676	0.268	1.702	−0.831	0.406
Age group: 60-69 years vs 18-39 years	0.731	0.298	1.791	−0.686	0.493
Age group: 70-79 years vs 18-39 years	0.815	0.336	1.978	−0.451	0.652
Age group: 80+ years vs 18-39 years	0.815	0.337	1.970	−0.454	0.650
Income Quintiles: 2 vs 1 (lowest)	1.017	0.823	1.257	0.156	0.876
Income Quintiles: 3 vs 1	0.987	0.792	1.229	−0.117	0.907
Income Quintiles: 4 vs 1	1.033	0.817	1.306	0.269	0.788
Income Quintiles: 5 (highest) vs 1	1.116	0.899	1.386	0.993	0.321
Cumulative low dose SGLT-2i medication[a]	1.010	0.592	1.723	0.036	0.971
Cumulative high dose SGLT-2i medication[a]	0.819	0.552	1.215	−0.993	0.321

[a]Cumulative dose measured with 90-days latency period at time t (per year).
[b]95% confidence interval.

5 | Discussion

In this article, we demonstrated the application of continuous-time marginal structural models with marked point process weights to assess the cumulative treatment effect of SGLT-2i prescriptions on the recurrent events of UTI in diabetes patient population with history of CKD. Unlike the earlier work of discrete-time marginal structural models [1], we considered time-dependent confounding with treatment-confounder feedback using the continuous-time treatment weights [11]. As an extension to earlier research [23], we employed marked point process to characterize the continuous-time treatment weights using visit process and dose assignment, and subsequently applied constrained optimization to improve the statistical properties of continuous-time treatment weights [32]. In a similar spirit to Santacatterina et al. [32], we implemented the constrained optimization of treatment weights using a prespecified precision parameter. We extended the earlier conceptualization of optimal treatment weights [29] in two different directions: (i) using recurrent events rather than survival events and (ii) using continuous-time stochastic process rather than discrete-time framework of marginal structural models. In longitudinal cohorts, the influence of time-varying confounder on the appropriate (causal) parameterization of how the treatment process affects the outcome process is not trivial in practice, especially in the presence of treatment-confounder feedback [53]. As an example, endowing the hazard ratio with causal interpretation is problematic because conditioning on recent survival may open noncausal pathways [10], and thereby introducing selection bias or inappropriately averaging the hazard ratio over time depending on the model parameterization [54]. Due to these shortcomings of hazard ratios, the causal interpretation is often characterized using Kaplan–Meier survival curves or cumulative incidence curves [10, 55].

We acknowledge some limitations of our approach, and how some assumptions may not necessarily hold in practice. First and foremost, the marked point process defined the dose process and the outcome process to be the bivariate marks, and the visit process to be the point process. A possible extension may include the construction of clustered marked point process in which a latent process influences both the visit process and the dose process [56]. Another avenue to explore is to incorporate the constrained optimization of treatment weights using

semiparametric estimation of structural failure time models formulated as continuous-time processes [57]. Future research may also focus on extending the continuous-time optimal weights to account for the time-invariant (subject-specific) latent variable between outcome process and treatment process. For example, the outcome process and the treatment process remain dependent conditional on the history through the unmeasured latent factor. The subject-specific latent variable may not allow the factorization of joint likelihoods with respect to the treatment and the outcome process. Heuristically, nonseparability of joint likelihood using two processes may produce similar random phenomena, and thus cannot be considered as independent. However, constrained optimization of continuous-time weights may provide an avenue to account for this subject-specific latency structure [58]. Lastly, we estimated the optimal treatment weights using the minimization of an Euclidean norm as the objective function (see Equation 11). Future extension of this work may consider an alternative to this objective function (including Bregman divergence, Kullback–Leibler divergence and Mahalanobis distance) to estimate the most optimal precision parameter using the criteria of bias-variance trade-off [59]. We empirically evaluated the validity of the stability assumption by fitting the conditional outcome model under the experimental measure and the observational measure. We noted that the stability assumption may not hold when the data generating mechanism parameterized the outcome with respect to cumulative duration of the treatment, and with the covariate intensity function. Alternative reparameterization of the outcome in the data generating mechanism might be necessary so that the stability assumption is not violated. Another option is to employ the pooled logistic regression with the basis-spline function to capture the time effect when stability assumption is not satisfied.

The empirical results of the simulation study demonstrated the application of continuous-time weights for recurrent events. The estimation of the treatment effect may exacerbate bias when the treatment or the outcome models are misspecified in the marginal structural model. In such scenarios, we may benefit from the application of doubly robust estimators [60], or alternative methods for estimating the optimal weights, including boosted trees or classification trees [61]. We applied a nonparametric counting process to characterize the intensity functions in the simulation study. A possible extension may include a parameterized counting process (e.g., Weibull distribution), which can be used to characterize the non-regularities in visit patterns, representing the instantaneous intensity of visit process with respect to the study follow-up [62, 63]. This extension may also share some commonality to earlier work in which a class of inverse-intensity rate-ratio weighted estimators were proposed using generalized estimating equations for outcome-dependent follow-up [9]. Apart from acknowledging the inherent complexities of causal methods in longitudinal cohorts, we need to acknowledge the computational challenges associated with the analysis of continuous-time marginal structural models when working with large data repositories. For example, the nested case control in cross-sectional studies, or its counterpart of case-base sampling in longitudinal cohorts may reduce the computational burden when working with large data repositories [64]. In this article, it was necessary to include the increased level of generality to illustrate the benefits of incorporating the continuous-time framework for the estimation of marginal structural models with marked point process weights [3, 6]. Future research may build on this methodological framework for several applications in other biomedical context including cardiovascular events for heart failure [22], and oncology events for prostate cancer [10].

6 | Conclusion

Our analyses did not support short-term or long-term increase in the incidence of UTI when low dose or high dose of SGLT-2i medications were prescribed. These findings supported the earlier scientific evidence in which no dose dependence of SGLT-2i medication was observed for several adverse clinical events including UTIs [46, 65, 66]. However, there exist earlier evidence on increased risk of genital fungal infections due to SGLT-2i medications [65, 67, 68]. Further research is required to verify the biological relationship between dose dependence of SGLT-2i medications and UTIs using appropriate emulation of randomized experiments from clinical data repositories elsewhere [16, 69].

Acknowledgments

This statistical research was supported by Natural Sciences and Engineering Research Council (NSERC) Canadian Graduate Scholarship (CGS: 534600).

Ethics Statement

This project received Research Ethics Board (REB) approval from Health Sciences REB board at University of Toronto (RIS Protocol Number: 39268).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Research data are not shared.

References

1. J. M. Robins, M. A. Hernan, and B. Brumback, "Marginal Structural Models and Causal Inference in Epidemiology," *Epidemiology* 11, no. 5 (2000): 550–560.

2. R. Sendor and T. Stürmer, "Core Concepts in Pharmacoepidemiology: Confounding by Indication and the Role of Active Comparators," *Pharmacoepidemiology and Drug Safety* 31, no. 3 (2022): 261–269.

3. Y. Xiao, M. Abrahamowicz, and E. E. M. Moodie, "Accuracy of Conventional and Marginal Structural Cox Model Estimators: A Simulation Study," *International Journal of Biostatistics* 6, no. 2 (2010): 13.

4. M. A. Hernán, B. Brumback, and J. M. Robins, "Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments," *Journal of the American Statistical Association* 96, no. 454 (2001): 440–448.

5. R. B. D'Agostino, M.-L. Lee, A. J. Belanger, L. A. Cupples, K. Anderson, and W. B. Kannel, "Relation of Pooled Logistic Regression to Time Dependent Cox Regression Analysis: The Framingham Heart Study," *Statistics in Medicine* 9, no. 12 (1990): 1501–1515.

6. S. Ferreira Guerra, M. E. Schnitzer, A. Forget, and L. Blais, "Impact of Discretization of the Timeline for Longitudinal Causal Inference Methods," *Statistics in Medicine* 39, no. 27 (2020): 4069–4085.

7. R. Colgan and M. Williams, "Diagnosis and Treatment of Acute Uncomplicated Cystitis," *American Family Physician* 84, no. 7 (2011): 771–776.

8. O. Saarela and Z. Liu, "A Flexible Parametric Approach for Estimating Continuous-Time Inverse Probability of Treatment and Censoring Weights," *Statistics in Medicine* 35, no. 23 (2016): 4238–4251.

9. P. Bužková and T. Lumley, "Semiparametric Modeling of Repeated Measurements Under Outcome-Dependent Follow-Up," *Statistics in Medicine* 28, no. 6 (2009): 987–1003.

10. P. C. Ryalen, M. J. Stensrud, S. Fosså, and K. Røysland, "Causal Inference in Continuous Time: An Example on Prostate Cancer Therapy," *Biostatistics* 21, no. 1 (2020): 172–185.

11. K. Røysland, "A Martingale Approach to Continuous-Time Marginal Structural Models," *Bernoulli* 17, no. 3 (2011): 895–915, https://doi.org/10.3150/10-BEJ303.

12. M. A. Hernán, M. McAdams, N. McGrath, E. Lanoy, and D. Costagliola, "Observation Plans in Longitudinal Studies With Time-Varying Treatments," *Statistical Methods in Medical Research* 18, no. 1 (2009): 27–52.

13. P. L. Spirtes, C. Meek, and T. S. Richardson, "Causal Inference in the Presence of Latent Variables and Selection Bias," arXiv preprint arXiv:1302.4983, (2013).

14. A. K. G. Jensen, H. Ravn, S. Sørup, and P. Andersen, "A Marginal Structural Model for Recurrent Events in the Presence of Time-Dependent Confounding: Non-specific Effects of Vaccines on Child Hospitalisations," *Statistics in Medicine* 35, no. 27 (2016): 5051–5069.

15. W. Hua, H. Mei, S. Zohar, M. Giral, and Y. Xu, "Personalized Dynamic Treatment Regimes in Continuous Time: A Bayesian Approach for Optimizing Clinical Decisions With Timing," *Bayesian Analysis* 17, no. 3 (2022): 849–878.

16. M. A. Hernán and J. M. Robins, "Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available," *American Journal of Epidemiology* 183, no. 8 (2016): 758–764.

17. J. H. Y. Wu, C. Foote, J. Blomster, et al., "Effects of Sodium-Glucose Cotransporter-2 Inhibitors on Cardiovascular Events, Death, and Major Safety Outcomes in Adults With Type 2 Diabetes: A Systematic Review and Meta-Analysis," *Lancet Diabetes and Endocrinology* 4, no. 5 (2016): 411–419.

18. R. G. Moses, S. Colagiuri, and C. Pollock, "Sglt2 Inhibitors: New Medicines for Addressing Unmet Needs in Type 2 Diabetes," *Australasian Medical Journal* 7, no. 10 (2014): 405.

19. D. Li, T. Wang, S. Shen, Z. Fang, Y. Dong, and H. Tang, "Urinary Tract and Genital Infections in Patients With Type 2 Diabetes Treated With Sodium-Glucose Co-Transporter 2 Inhibitors: A Meta-Analysis of Randomized Controlled Trials," *Diabetes, Obesity and Metabolism* 19, no. 3 (2017): 348–355.

20. C. V. Dave, S. Schneeweiss, and E. Patorno, "Comparative Risk of Genital Infections Associated With Sodium-Glucose Co-Transporter-2 Inhibitors," *Diabetes, Obesity and Metabolism* 21, no. 2 (2019): 434–438.

21. R. J. Cook and J. F. Lawless, *The Statistical Analysis of Recurrent Events* (Springer, 2007).

22. L. Mancini and A. M. Paganoni, "Marked Point Process Models for the Admissions of Heart Failure Patients," *Statistical Analysis and Data Mining: The ASA Data Science Journal* 12, no. 2 (2019): 125–135.

23. Y. Dong, *Continuous-Time Marginal Structural Models for Adverse Drug Effects in Pharmacoepidemiology* (University of Toronto (Canada), 2021).

24. A. P. Dawid and V. Didelez, "Identifying the Consequences of Dynamic Treatment Strategies: A Decision-Theoretic Overview," *Statistics Surveys* 4 (2010): 184–231.

25. L. Meira-Machado, J. de Uña-Álvarez, C. Cadarso-Suárez, and P. K. Andersen, "Multi-State Models for the Analysis of Time-To-Event Data," *Statistical Methods in Medical Research* 18, no. 2 (2009): 195–222.

26. P. K. Andersen and R. D. Gill, "Cox's Regression Model for Counting Processes: A Large Sample Study," *Annals of Statistics* 10, no. 4 (1982): 1100–1120, https://doi.org/10.1214/aos/1176345976.

27. L. D. A. F. Amorim and J. Cai, "Modelling Recurrent Events: A Tutorial for Analysis in Epidemiology," *International Journal of Epidemiology* 44, no. 1 (2015): 324–333, https://doi.org/10.1093/ije/dyu222.

28. O. Aalen, O. Borgan, and H. Gjessing, *Survival and Event History Analysis: A Process Point of View* (Springer Science & Business Media, 2008).

29. M. Santacatterina and M. Bottai, "Optimal Probability Weights for Inference With Constrained Precision," *Journal of the American Statistical Association* 113, no. 523 (2018): 983–991.

30. E. M. Pullenayegum and B. M. Feldman, "Doubly Robust Estimation, Optimally Truncated Inverse-Intensity Weighting and Increment-Based Methods for the Analysis of Irregularly Observed Longitudinal Data," *Statistics in Medicine* 32, no. 6 (2013): 1054–1072.

31. O. Bembom and M. J. van der Laan, *Data-Adaptive Selection of the Truncation Level for Inverse-Probability-of-Treatment-Weighted Estimators*. Division of Biostatistics Working Paper Series. Working Paper 230 (*U.C. Berkeley*, 2008).

32. M. Santacatterina, C. García-Pareja, R. Bellocco, A. Sönnerborg, A. M. Ekström, and M. Bottai, "Optimal Probability Weights for Estimating Causal Effects of Time-Varying Treatments With Marginal Structural Cox Models," *Statistics in Medicine* 38, no. 10 (2019): 1891–1902.

33. J. R. Zubizarreta, "Stable Weights That Balance Covariates for Estimation With Incomplete Outcome Data," *Journal of the American Statistical Association* 110, no. 511 (2015): 910–922.

34. M. A. Hernán and J. M. Robins, *Causal Inference: What If* (Chapman & Hall/CRC, 2024).

35. B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding," *Journal of Optimization Theory and Applications* 169, no. 3 (2016): 1042–1068.

36. A. Burton, D. G. Altman, P. Royston, and R. L. Holder, "The Design of Simulation Studies in Medical Statistics," *Statistics in Medicine* 25, no. 24 (2006): 4279–4292.

37. K. Tu, B. Aliarzadeh, T. Chen, and S. Kalia, "University of Toronto Family Medicine Report: Caring for Our Diverse Population," (2020), https://issuu.com/dfcm/docs/technical_appendix_final_27oct2020.

38. T. Williamson, M. E. Green, R. Birtwhistle, et al., "Validating the 8 Cpcssn Case Definitions for Chronic Disease Surveillance in a Primary Care Database of Electronic Health Records," *Annals of Family Medicine* 12, no. 4 (2014): 367–372.

39. K. Yau, A. Dharia, I. Alrowiyti, and D. Z. I. Cherney, "Prescribing sglt2 Inhibitors in Patients With Ckd: Expanding Indications and Practical Considerations," *Kidney International Reports* 7, no. 7 (2022): 1463.

40. K. J. Rothman, S. Greenland, T. L. Lash, et al., *Modern epidemiology*, vol. 3 (Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008).

41. A. V. Rigobon, S. Kalia, J. Nichols, et al., "Impact of the Diabetes Canada Guideline Dissemination Strategy on the Prescription of Vascular Protective Medications: A Retrospective Cohort Study, 2010–2015," *Diabetes Care* 42, no. 1 (2019): 148–156.

42. H. Jeremiah, L. Sung, J. C. Kwong, R. Sutradhar, K. Tu, and J. D. Pole, "Use of Physician Billing Claims to Identify Infections in Children," *PLoS One* 13, no. 11 (2018): e0207468.

43. N. Hammar, B. Farahmand, M. Gran, S. Joelson, and S. W. Andersson, "Incidence of Urinary Tract Infection in Patients With Type 2 Diabetes. Experience From Adverse Event Reporting in Clinical Trials," *Pharmacoepidemiology and Drug Safety* 19, no. 12 (2010): 1287–1292.

44. M. Greiver, A. Havard, J. K. F. Bowles, et al., "Trends in Diabetes Medication Use in Australia, Canada, England, and Scotland: A Repeated Cross-Sectional Analysis in Primary Care," *British Journal of General Practice* 71, no. 704 (2021): e209–e218.

45. D. Fitchett, "A Safety Update on Sodium Glucose Co-Transporter 2 Inhibitors," *Diabetes, Obesity and Metabolism* 21 (2019): 34–42.

46. F.-H. Shi, H. Li, J. Yue, et al., "Clinical Adverse Events of High-Dose vs Low-Dose Sodium–Glucose Cotransporter 2 Inhibitors in Type 2 Diabetes: A Meta-Analysis of 51 Randomized Clinical Trials," *Journal of Clinical Endocrinology and Metabolism* 105, no. 11 (2020): 3600–3611.

47. CPCSSN, "Cpcssn Data Dictionary, Version 4.0.3," (2020), http://cpcssn.ca/wp-content/uploads/2018/04/CPCSSN-Data-Dictionary-v4.0.3.pdf.

48. S. Kalia, O. Saarela, M. Escobar, R. Moineddin, and M. Greiver, "Estimation of Marginal Structural Models Under Irregular Visits and Unmeasured Confounder: Calibrated Inverse Probability Weights," *BMC Medical Research Methodology* 23, no. 1 (2023): 1–16.

49. S. Kalia, O. Saarela, T. Chen, et al., "Marginal Structural Models Using Calibrated Weights With Superlearner: Application to Type Ii Diabetes Cohort," *IEEE Journal of Biomedical and Health Informatics* 26, no. 8 (2022): 4197–4206.

50. S. Kalia, O. Saarela, B. O'Neill, et al., "Emulating a Target Trial Using Primary-Care Electronic Health Records: Sodium-Glucose Cotransporter 2 Inhibitor Medications and Hemoglobin a1c," *American Journal of Epidemiology* 192, no. 5 (2023): 782–789.

51. CPCSSN, "Case Definitions: Canadian Primary Care Sentinel Surveillance Network (cpcssn), Version 2020-q2, (May 15, 2020)." http://cpcssn.ca/wp-content/uploads/2020/10/CPCSSN-Case-Definitions-2020-Q2-1.pdf.

52. V. Didelez, "Graphical Models for Marked Point Processes Based on Local Independence," *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 70, no. 1 (2008): 245–264.

53. O. O. Aalen, J. M. Gran, K. Røysland, M. J. Stensrud, and S. Strohmaier, "Feedback and Mediation in Causal Inference Illustrated by Stochastic Process Models," *Scandinavian Journal of Statistics* 45, no. 1 (2018): 62–86.

54. M. A. Hernán, "The Hazards of Hazard Ratios," *Epidemiology* 21, no. 1 (2010): 13–15.

55. E. E. M. Moodie, D. A. Stephens, and M. B. Klein, "A Marginal Structural Model for Multiple-Outcome Survival Data: Assessing the Impact of Injection Drug Use on Several Causes of Death in the Canadian Co-Infection Cohort," *Statistics in Medicine* 33, no. 8 (2014): 1409–1425.

56. T. Zhang, "On Independence and Separability Between Points and Marks of Marked Point Processes," *Statistica Sinica* 27, no. 1 (2017): 207–227.

57. S. Yang, K. Pieper, and F. Cools, "Semiparametric Estimation of Structural Failure Time Models in Continuous-Time Processes," *Biometrika* 107, no. 1 (2020): 123–136.

58. S. Yang, "Propensity Score Weighting for Causal Inference With Clustered Data," *Journal of Causal Inference* 6, no. 2 (2018): 20170027, https://doi.org/10.1515/jci-2017-0027.

59. L. M. Bregman, "The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming," *USSR Computational Mathematics and Mathematical Physics* 7, no. 3 (1967): 200–217.

60. J. D. Y. Kang, J. L. Schafer, and J. D. Y. Kang, "Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean From Incomplete Data," *Statistical Science* 22, no. 4 (2007): 523–539, https://doi.org/10.1214/07-sts227.

61. M. E. Karim, J. Petkau, P. Gustafson, H. Tremlett, and The Beams Study Group, "On the Application of Statistical Learning Approaches to Construct Inverse Probability Weights in Marginal Structural Cox Models: Hedging Against Weight-Model Misspecification," *Communications in Statistics - Simulation and Computation* 46, no. 10 (2017): 7668–7697.

62. R. J. Cook and J. F. Lawless, *Multistate Models for the Analysis of Life History Data* (Chapman and Hall/CRC, 2018).

63. C. C. T. Fok, J. O. Ramsay, M. Abrahamowicz, and P. Fortin, "A Functional Marked Point Process Model for Lupus Data," *Canadian Journal of Statistics* 40, no. 3 (2012): 517–529.

64. J. A. Hanley and O. S. Miettinen, "Fitting Smooth-In-Time Prognostic Risk Functions via Logistic Regression," *International Journal of Biostatistics* 5, no. 1 (2009), https://doi.org/10.2202/1557-4679.1125.

65. J. Wilding, "Sglt2 Inhibitors and Urinary Tract Infections," *Nature Reviews Endocrinology* 15, no. 12 (2019): 687–688.

66. W. Alkabbani, A. Zongo, J. K. Minhas-Sandhu, et al., "Sodium-Glucose Cotransporter-2 Inhibitors and Urinary Tract Infections: A Propensity Score–Matched Population-Based Cohort Study," *Canadian Journal of Diabetes* 46, no. 4 (2022): 392–403.

67. P. K. Bundhun, G. Janoo, and F. Huang, "Adverse Drug Events Observed in Patients With Type 2 Diabetes Mellitus Treated With 100 Mg Versus 300 Mg Canagliflozin: A Systematic Review and Meta-Analysis of Published Randomized Controlled Trials," *BMC Pharmacology and Toxicology* 18, no. 1 (2017): 1–12.

68. X. Dai, Z.-c. Luo, L. Zhai, W.-p. Zhao, and F. Huang, "Adverse Drug Events Associated With Low-Dose (10 Mg) Versus High-Dose (25 Mg) Empagliflozin in Patients Treated for Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials," *Diabetes Therapy* 9, no. 2 (2018): 753–770.

69. C. V. Dave, S. Schneeweiss, D. Kim, M. Fralick, A. Tong, and E. Patorno, "Sodium–Glucose Cotransporter-2 Inhibitors and the Risk for Severe Urinary Tract Infections: A Population-Based Cohort Study," *Annals of Internal Medicine* 171, no. 4 (2019): 248–256.