Bayesian analysis of longitudinal studies with treatment by indication

Reagan Mozer1 · Mark E. Glickman2

Received: 19 January 2022 / Accepted: 20 November 2022 / Published online: 15 December 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
In a medical setting, observational studies commonly involve patients who initiate a particular treatment (e.g., medication therapy) and others who do not, and the goal is to draw causal inferences about the effect of treatment on a time-to-event outcome. A difficulty with such studies is that the notion of a treatment initiation time is not well-defined for the control group. In this paper, we propose a Bayesian approach to estimate treatment effects in longitudinal observational studies where treatment is given by indication and thereby the exact timing of treatment is only observed for treated units. We present a framework for conceptualizing an underlying randomized experiment in this setting based on separating the time of indication for treatment, which we model using a latent state-space process, from the mechanism that determines assignment to treatment versus control. Next, we develop a two-step inferential approach that uses Markov Chain Monte Carlo (MCMC) posterior sampling to (1) infer the unobserved indication times for units in the control group, and (2) estimate treatment effects based on inferential conclusions from Step 1. This approach allows us to incorporate uncertainty about the unobserved indication times which induces uncertainty in both the selection of the control group and the measurement of time-to-event outcomes for these controls. We demonstrate our approach to study the effects on mortality of inappropriately prescribing phosphodiesterase type 5 inhibitors (PDE5Is), a medication contraindicated for certain types of pulmonary hypertension, using data from the Veterans Affairs (VA) health care system.

Keywords Causal inference · Observational studies · Comparative effectiveness research

1 Introduction
In certain observational studies, particularly in a medical context, interest centers on estimating the causal effects of an action, treatment, or intervention on a time-to-event outcome (e.g., the effects of surgery on post-operative survival time). Insights about the
effects of these non-randomized treatments have the potential to help answer important questions both in population health and in individualized medicine. In these settings, measuring time-to-event outcomes in a treated group is straightforward; time is measured from treatment initiation. However, a key obstacle is that time-to-event outcomes in the control group are not defined because the notion of an initiation time for controls is an elusive construct. The time at which a study participant “initiates control” is difficult to conceptualize. The lack of initiation times in the control group therefore presents substantial challenges in making causal inferences from time-to-event outcomes.

The question of how to make valid causal inferences in this setting using longitudinal, observational data becomes further complicated with the possibility of confounding due to “treatment by indication” (Poses et al. 1995). This issue stems from the principle that a good clinician will initiate a new treatment only when evidence exists that the candidate treatment is medically beneficial or necessary for the patient in question (Poses et al. 1995). In studies evaluating the effects of a non-randomized treatment on a population with a common disease or condition, this means that treatment is initiated (and, therefore, assignment to treatment is observed) only when the potential benefits outweigh the anticipated risks or side effects associated with treatment. As a result, subjects who receive treatment during the study period may differ systematically from untreated subjects observed during the same time period in terms of their prognosis.

Confounding due to treatment by indication can generally be avoided by comparing groups of subjects who presented with similar indications at similar times after their initial diagnosis. That is, patients who ideally express similar symptoms and have similar health characteristics indicating the necessity of medical intervention at the same time after disease diagnosis can be viewed as a cohort within which the causal effect of treatment can be measured. However, these variables are typically unrecorded for subjects who do not receive treatment. Thus, for the group of untreated subjects, it is unclear whether medical intervention was indicated at an unknown time during the study and treatment was purposefully withheld at that time or if treatment was simply never considered due to lack of indications. The former could happen, for example, if a patient’s health had deteriorated from a disease, and the provider considered prescribing a specific medication (the treatment) or considered withholding the medication in favor of another form of therapy (the control) that might not have been recorded. This issue is especially common in studies that rely on data from electronic medical records, administrative databases, and health registries (Byar 1980; Levine and Julian 2008). In these settings, unknown assignment to treatment or control may be missing due to measurement errors (e.g., if receipt of the control intervention is not recorded) or may be (right-) censored due to follow-up (e.g., for patients whose assignment to treatment or control occurs after the study has ended). When criteria for determining indications for treatment cannot be applied to the pool of potential controls, a possible strategy is to infer indication times for these units. A primary challenge in this domain therefore concerns how to infer the unknown indication times from the observed data.

In this paper, we consider how to draw causal inferences about a binary treatment, which can either be initiated or withheld from a given subject at a single point in time. This time point, which we refer to as the “indication time”, is the time at which a subject presents with a particular set of symptoms or pre-specified indications that trigger clinical intervention. In a medical setting, the indication time for any individual might occur when a subject needs medication therapy, or might be solely determined by clinical factors. For instance, in our applied example, we consider a population of patients in the VA health system who are diagnosed with certain types of pulmonary hypertension (PH), where individuals may
receive medication therapy in the form of phosphodiesterase-5-inhibitors (PDE5Is) for treatment of PH-related symptoms only when it is determined that the medication may be beneficial given the severity of the patient’s condition. As a result, indication times may vary greatly across patients in a sample, which induces a complicated time structure in the observed data. To accommodate this structure, we propose an approach for treatment effect estimation that views indication times as a fixed but possibly unknown pre-treatment covariate representing relevant characteristics related to the patient’s health. We then condition on these indication times to construct an estimator of the causal effect of treatment versus control that is independent from the time-structure of the underlying data.

The paper proceeds as follows. In Sect. 2, we first review existing methods in this domain and describe a general strategy for making causal inferences in longitudinal observational studies by approximating a sequence of randomized trials. We then present an alternative conceptualization of the data-generating process that can facilitate causal inference in this setting, which is based on separating the process that governs the time of indication for treatment from the mechanism that determines the receipt of treatment versus control. In Sect. 3, we present a framework for designing longitudinal observational studies to approximate a hypothetical randomized experiment based on this conceptualization and describe the formal assumptions required for inference. The core of this framework is a joint state-space model for predicting indication times as a function of longitudinal covariates, described in Sect. 4, which we use to infer the missing indication times for untreated patients. In this section, we also propose an approach for estimating treatment effects that directly incorporates uncertainty about the inferred indication times. We then illustrate the proposed framework in Sect. 5 to study the effects of prescribing contraindicated PDE5Is for treatment of PH using data from the VA health care system.

## 2 Background

When attempting to draw causal inferences in observational studies, it may be desirable to “design” the study in a manner that approximates a randomized experiment (Rubin 1984). While this design-based approach to covariate adjustment in observational studies has been widely applied to estimate causal effects using cross-sectional data, little work has focused on extending these methods to longitudinal settings. In fact, how to properly design and analyze longitudinal studies of non-randomized medical interventions (comparative effectiveness studies) remains a point of controversy (Rubin 2010). One strategy that is commonly used in observational studies is to select a set of “not-yet-treated” controls—units who do not receive treatment during the study period—based on their similarity to treated units with respect to pre-treatment variables observed during a baseline period. However, control group selection becomes complicated in longitudinal settings where units may enter the study at different points in calendar time. Instead of having a single baseline period for all units, in this setting we observe a (potentially unique) baseline period for each treated unit, defined as the time from study entry to the time of treatment initiation. The primary challenge is then how to define a comparable pre-treatment period for potential controls who do not receive treatment at any point during the study window (Basse et al. 2016).

To address this issue, some studies have attempted to identify alternative indication times (i.e., “phantom” treatment times) for untreated units for whom there is no receipt of treatment. For instance, in an observational study evaluating the effects
of selective cyclo-oxygenase-2 inhibitors for treatment of upper gastrointestinal hemorrhage in elderly patients, Mamdani et al. (2002) randomly assigned indication times for untreated units. Zhou et al. (2005) proposed a similar approach for prescription time-distribution matching, whereby indication times for each untreated unit are selected at random from the distribution of indication times observed among the treated units. This strategy ensures that the treatment and control groups will be balanced with respect to time of treatment initiation, but requires the strong assumption that indication times are independent and identically distributed for all units in the study population.

Some studies have addressed this issue by identifying a proxy for the time of treatment indication (e.g., dispensement dates of other prescription medications during the study period) (McGettigan and Henry 2006). This approach is the basis of the “active-comparator design,” which restricts the control group to the set of untreated units who received another active drug during the study period (Yoshida et al. 2015). The core idea of this design is that any given sample of untreated units is likely to include subjects with no indications for treatment (e.g., mild disease) as well as subjects with contraindications for the treatment of interest (e.g., due to comorbid conditions). While this framework provides a useful template for designing comparative effectiveness studies, if the treatment of interest is the most commonly used therapy for a given condition (i.e., a first-line therapy) and the alternatives are used infrequently, selecting an active comparator can be challenging.

Other methods have attempted to address these issues by framing the data-generating process in terms of a sequence of randomized experiments occurring over time, where at each time the treatment can either be initiated or withheld (Robins et al. 2000; Hernán et al. 2008). One such strategy is so-called risk set matching (RSM; Li et al. 2001), which aims to assess the effects of different treatment sequences on an outcome observed at a fixed end-point, thereby avoiding the issue of control group identification. This strategy makes the implicit assumption that, unlike in a classical randomized experiment, all patients will receive the treatment of interest eventually and it is rather their times of treatment that are a result of randomization (van Houwelingen and Putter 2011). Accordingly, causal effects are defined at each of a set of discrete time points and can be estimated by comparing units who received treatment at that time with units who were not yet treated at that point. The resulting inferences may be useful to practitioners interested in estimating the effects of delaying a particular treatment and may help answer questions about the optimal time to apply a medical intervention for patients that require treatment (Danaei et al. 2013; Watson et al. 2019). However, inferences based on RSM may not be appropriate in settings with treatment by indication, where the time of initiation is not under the control of the treatment initiator, usually the clinician. When a patient presents with symptoms that indicate the need for medical intervention, their health care provider is faced with a decision about which among available treatments to initiate at that time. In these settings, it is not sensible to hypothesize about how a patient’s outcomes would change if their indications had presented at a different time (Angrist et al. 1996). Rather, the time of indication for treatment can be viewed as a fixed pre-treatment covariate (i.e., a variable that is unaffected by assignment to treatment or control) that characterizes the severity of the underlying condition, and causal effects should be framed as contrasts of the outcomes.
that would be observed for each unit under different treatment strategies initiated at that time.

3 A causal effects framework

The framework we propose mimics the protocol typically used to analyze data from a randomized controlled trial, where outcomes are observed during a pre-specified follow-up period beginning at the time of randomization, which is precisely the same time as initiation of therapy. To motivate this framework, consider the following hypothetical randomized experiment with a binary treatment, where interest is in estimating the causal effects of treatment on a time-to-event outcome for a particular patient population. Here, treatment is defined as the decision to apply the intervention, and control is defined as the decision to withhold the intervention. Assume that patients become eligible for inclusion in the experiment only when their health reaches a point where clinical intervention may be beneficial to the patient despite any associated risks or side effects (the so-called “indication time”). Upon enrollment, suppose that patients are then randomized to receive either treatment or control with probability that depends only on their indication time. After randomization, suppose each patient is observed until the earliest occurrence of a pre-specified event (e.g., renal failure) or the end of follow-up, whichever occurs first. The primary outcome is based on this event time calculated with respect to indication time. In this idealized experiment, contrasts of outcomes between treated and control units with similar indication times will be unbiased estimates of the causal effects of interest.

This conceptualization of the underlying randomized experiment explicitly defines the control group as the set of untreated patients for whom treatment was indicated during the study period and deliberately withheld (“true controls”). The remaining untreated patients are those for whom no indications for treatment were present during the study (“ineligible controls”). Because these patients are not assigned to treatment or control during the study period, they are not relevant units for the purposes of causal inference. In a setting in which indication times are fully observed, the set of ineligible controls can easily be identified and discarded prior to analysis. However, in many observational studies, indication times may be censored due to follow-up or death and are often at best partially observed, specifically, for only those patients receiving treatment. In addition, the time-varying probabilities of assignment to treatment are generally unknown and may be difficult to infer without expert domain knowledge.

3.1 Problem setup

We assume a cohort of \( N \) patients (i.e., observational study units), each of whom is observed over a specified study time window divided into \( K \) discrete time periods. If a patient in the cohort receives the treatment of interest during their observation period, then he or she is included in the study as a treated unit. Other members of the cohort are eligible to be controls during their respective observation periods. For each unit \( i = 1, \ldots, N \), suppose we observe a vector of \( p \) covariates collected at each of the \( K \) time periods, denoted \( X_i = (X_{i0}, \ldots, X_{iK}) \), where \( X_{i0} \) is a \( p \)-vector of baseline covariates.
observed at study entry. These covariates, which are assumed to be fully observed, capture characteristics of each unit or the unit’s environment (such as age, gender, physiological factors, diet, medical treatments, and environmental exposures) over the course of the study. Let $T_i$ denote the indication time period for unit $i$ relative to time of entry into the study, which may occur at a discrete time within the study or may be right-censored at the end of follow-up. By construction, units with $T_i > K$ do not receive either treatment or control during the study period and we let $S_i = 1 \{ T_i \in [0, K] \}$ be an indicator of eligibility for inclusion in the study as a control. Let $Z_i$ be an indicator for assignment to treatment upon indication for unit $i$, where

$$Z_i = \begin{cases} 0 & \text{if unit } i \text{ is assigned to control} \\ 1 & \text{if unit } i \text{ is assigned to treatment.} \end{cases}$$

In our setting, $T_i$ is typically observed only for units who are assigned to treatment and that the assignment takes place within the study period (i.e., when $Z_i = 1$ and $T_i \in [0, K]$). Finally, let $M_i = 1 - Z_i S_i$ be an indicator for the missingness of $T_i$ such that $M_i = 1$ for units whose indication times are missing due to either censoring (i.e., if $S_i = 0$) or treatment assignment (i.e., $Z_i = 0$).

Interest is in estimating the causal effects of assignment to treatment versus control on time-to-event outcomes defined as the time from indication for treatment to the first occurrence of an event of interest (e.g., death, hospital discharge, symptom remission). Under the Rubin Causal Model (RCM; Holland 1986), each participant has two potential outcomes, $Y_i(0)$ and $Y_i(1)$, which here represent the event times that would be observed for patient $i$ under assignment to control or treatment, respectively, upon indication to treatment. Due to the fundamental problem of causal inference, we can observe at most one of these potential outcomes (i.e., the potential outcome corresponding to the treatment actually received) for each subject in the study (Rubin 1976). Therefore, let $Y_{i, \text{obs}} = (1 - Z_i) Y_i(0) + Z_i Y_i(1)$ denote the observed outcome for patient $i$ under a given treatment assignment $Z_i$.

Table 1 shows the structure of the data in this setting. Here, it is important to distinguish between two distinct missing data mechanisms that give rise to the observed and missing values. The first type of missing data are the indication times that are naturally right-censored at the end of the study. Although, in principle, these values are observable over a sufficient follow-up period, the missing-data mechanism is determined by the specified observation period. The second type of missing data in this setting are the potential outcomes that are not possible to observe under a given treatment assignment. These missing potential outcomes are therefore “endogenous” missing values in the sense that the missingness mechanism is completely determined by treatment.

| Units | $Z_i \in \{0, 1\}$ | $T_i \in \{1, 2, \ldots\}$ | $X_i = (X_{i1}, \ldots, X_{iT_i})$ | $Y_i(0) - T_i$ | $Y_i(1) - T_i$ |
|-------|-----------------|-----------------|-----------------|----------------|----------------|
| 1     | 1               | $t_1$           | $x_1$           | $\star$        | $y_1 - t_1$   |
| 2     | 1               | $t_2$           | $x_2$           | $\star$        | $y_2 - t_2$   |
| 3     | ?               | ?               | ?               | $y_3 - ?$      | $\star$       |
| 4     | ?               | ?               | ?               | $y_4 - ?$      | $\star$       |

$\star$ denotes endogenous missingness and ‘?’ denotes exogenous missingness.
assignment. In comparison, the missing indication times $T_i$ among untreated units are “exogenous” missing values, since their missingness depends, in part, on external factors such as the duration of the study period. As in a standard randomized experiment, the unit-level missing potential outcomes are impossible to observe under a given treatment assignment, and the goal of causal inference is therefore to recover these values under plausible modeling assumptions in order to make inferences about the causal effects of interest.

### 3.2 Causal estimands and assumptions for identification

For the setting described above, the unit-level treatment effect for a patient with fixed indication time $T_i = t$ is defined as

$$\tau_i = (Y_i(1) - t) - (Y_i(0) - t) = Y_i(1) - Y_i(0).$$

Under the Stable Unit Treatment Value Assumption (SUTVA; Rubin 1980) which asserts that there is no interference between units and no hidden forms of treatment, the conditional average treatment effect (CATE) for patients with indication time $T_i = t$ is then

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|T_i = t],$$

(1)

where $\mathbb{E}[\cdot]$ is the average across eligible units indexed by $i = 1, \ldots, N$. In longitudinal settings where the outcome of interest is defined relative to a time of death or failure, $\tau$ captures the average change in survival time under treatment compared to control for units who present with indications for treatment at particular time points $t = 1, \ldots, K$ during the study period. We can also construct an aggregate measure of these time-specific effects as

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|T_i \in [0, K]] = \frac{1}{K} \sum_{t=1}^{K} \tau_t,$$

(2)

which captures the average treatment effect (ATE) for eligible units who have indications for treatment at any point during follow-up.

To construct an unbiased estimator of $\tau_t$ for each $t = 1, \ldots, K$ using non-randomized data, we assume that assignment to treatment versus control upon indication is conditionally independent of the potential outcomes given all pre-treatment covariates including the indication times (i.e., $(Y_i(0), Y_i(1)|Z_i, T_i$ for all $i = 1, \ldots, N$). Given a particular randomization $Z$ and indication times $T$, we can then construct a simple estimator for (1) through straightforward contrasts of the observed potential outcomes within each treatment group (e.g., $\hat{\tau} = \mathbb{E}[Y_{i, \text{obs}}|Z_i = 1, T_i = t] - \mathbb{E}[Y_{i, \text{obs}}|Z_i = 0, T_i = t]$). Here, in contrast to the classical setting, the outcomes that are actually observed for each unit depend not only on their treatment assignment but also on their indication time. For units whose indication times are not observed, these missing times must be inferred in order to evaluate the distribution of potential outcomes. Note that with an untreated comparison group, estimating the average treatment effect for the treated (ATT) without conditioning on the times of indication may be preferred, since we would rarely want to initiate treatment for all units in the population. However, in the setting of treatment by indication or when an active comparator is present, estimating the average treatment effect for the entire population is desirable because we are focusing on the choice of one treatment compared to another assuming that indications for treatment have already occurred.
3.3 Overview of the inferential approach

To infer the missing indication times among potential controls, the first goal is to construct a model for predicting the observed indication times $T_{\text{obs}}$ based on baseline and time-varying covariates $X_i$. Inference from this model will also induce inferences on the potential outcomes in the control group, since the observed outcomes must be calculated relative to the indication times. By applying the assumptions described above in Sect. 3.2, we can factor the joint distribution of the complete data, including all observed potential outcomes $Y_{T_{\text{obs}}}$, as well as both the observed and unobserved indication times, $T = (T_{\text{obs}}, T_{\text{mis}})$, given some global parameter $\Theta$ partitioned as $\Theta = (\theta_1, \theta_2, \theta_3)$ as

$$p(Y_{T_{\text{obs}}}, T, Z|X, \Theta) = p(Y_{T_{\text{obs}}}|T, Z, X, \theta_1)p(Z|T, X, \theta_2)p(T|X, \theta_3).$$

For Bayesian inference with prior density $p(\theta_1, \theta_2, \theta_3)$, the posterior density of $\Theta$ given the complete data is

$$p(\Theta|Y_{T_{\text{obs}}}, T, Z, X) \propto p(Y_{T_{\text{obs}}}|T, Z, X, \theta_1)p(Z|T, X, \theta_2)p(T|X, \theta_3)p(\theta_1, \theta_2, \theta_3).$$

Under this framework, the prior $p(\theta_1, \theta_2, \theta_3)$ may be customized to the application at hand. Given the distinct role of the parameters, we would expect the joint prior to factor into a product of three independent densities. Thus, if prior information exists about one or more of the parameters as part of its model component (e.g., if one expects certain covariates to be more or less influential in determining the time of treatment indication), then this information can be directly incorporated into the model via an informative prior density. Otherwise, if no such prior information is available, our framework also permits non-informative prior components. Posterior inference on $\theta$ accounting for the missing values can then proceed by straightforward application of Markov Chain Monte Carlo (MCMC) techniques, such as the Gibbs sampler (Geman and Geman 1984; Gelman et al. 2014). See Appendix A for an example of one such posterior sampling scheme.

3.4 A simulation study

To demonstrate our inferential procedure on a basic version of our modeling framework, we conducted a simulation study in which we evaluated the proposed method for different sample sizes and different amounts of missing data. Using simulated data, we explored how our framework can recover the parameters of the treatment assignment and indication time models (i.e., $p(Z|T, X, \theta_2)$ and $p(T|X, \theta_3)$) in a simple data-generating scenario, assuming the model has been correctly specified. We note that, because our proposed framework involves addressing the parameters of the outcome model, $p(Y_{T_{\text{obs}}}|T, Z, X, \theta_1)$, as a separate inferential step, our simulations are conducted without access to outcome data.

For simplicity, we used a fixed study period of $K = 3$ time intervals and a single time-varying covariate $X = (X_1, X_2, X_3)$. Covariate values at each time interval were simulated for each subject from a Bernoulli distribution with fixed parameter $\pi_X$. Given $X$, indication times were generated for each subject using the following model:

$$p(T = 1) = q_1^{X_1}q_2^{(1-X_1)}$$
$$p(T = t|T > t - 1) = q_1^{X_t}q_2^{(1-X_t)} \text{ for } t = 2, 3,$$
where \( q_1 \) and \( q_2 \) are probability parameters. For the purposes of these simulations, we assumed that treatment assignment is conditionally independent of \( X \) given indication times \( T \). For each subject, we simulated treatment assignment from a Bernoulli distribution with time-specific probability parameter \( p_t = P(Z = 1 | T = t) \). Finally, we determined the missingness indicator \( M \), which takes value 1 for subjects with \( T > K \) or \( Z = 0 \), and 0 for subjects with \( T \leq K \) and \( Z = 1 \).

We conducted a set of simulations in which we varied the sample size \( n \) and covariate probability \( \pi_X \). In each scenario, we generated 1000 data sets and fit the model defined above using MCMC posterior simulation with Uniform priors on all model parameters. Each MCMC sampler was run with two parallel chains using a burn-in period of 5,000 draws and ran for between 10,000 to 350,000 iterations until convergence, which was verified using the Geweke (1992) and Gelman and Rubin (1992) statistics applied to the model parameters. We also used plots to assess the extent of the autocorrelation of the samplers. For each simulation condition, the average parameter estimate was calculated by averaging the median of the posterior distribution of a parameter across data sets in a simulation condition. We note that the mean of the posterior distribution for each parameter resulted in very similar parameter estimates.

Table 2 summarizes the simulation results for all model parameters. As can be seen in the table, our inferential procedure yielded acceptable coverage rates for all model parameters under nearly every simulation condition. We observe slight undercoverage when the covariate probability was \( \pi_X = 0.3 \) and the sample size was relatively large (\( n = 500 \) or \( n = 1000 \)), but given the non-monotonic patterns in coverage with respect to values of \( \pi_X \) this may be a consequence of Monte Carlo error. In addition, we tended to slightly underestimate the parameters related to indication time (i.e., \( \theta_1 = (q_1, q_2) \)) while slightly overestimating the parameters that determine treatment assignment at the time of indication (i.e., \( \theta_2 = (p_1, p_2, p_3) \)). However, these biases generally decreased as the sample size increased.

### 4 State-space model for time of treatment indication

Based on the conceptual framework presented in Sect. 3, we propose a specific and pragmatic model for predicting the time of indication for treatment - the earliest time at which a patient presents with indications for clinical intervention - as a function of both fixed and time-varying covariates. In particular, we hypothesize that observed covariate measurements that reflect worsening health and diminished functional capacity will be predictive of indication times. Similarly, we assume that covariates capturing provider characteristics or temporal features (e.g., month or year when measurements are recorded) are independent of the indication times but may influence the probability of assignment to treatment upon indication. For instance, in the application presented in Sect. 5, we expect the probability of treatment to decline over time as clinicians become more informed about populations where the medication of interest may be contraindicated. We capture these dependencies through separate components of a hierarchical model; the first component characterizes the stochastic process that governs patients’ indication times, and the second component describes the conditional probability of assignment to treatment given the indication times.
4.1 Model formulation

The first component of our model adopts a so-called “threshold approach,” which views indication times as the first hitting times of a latent process (Albert and Chib 1993). In particular, suppose that the time series of covariate measurements for each unit, $X_i$, is independent from the measurements of other units, and let $\theta_i = (\theta_{i1}, \ldots, \theta_{iK})$ be a state variable representing random (i.e., unexplained) fluctuations in unit $i$’s health over the course of the study. We assume a first order autoregressive process for the daily health fluctuations for unit $i$ such that between time periods $t$ and $t - 1$.

$$\theta_{it} = \rho \theta_{i(t-1)} + e_{it}, \quad e_{it} \sim N(0, \sigma_e^2),$$

where $N$ denotes a normal distribution. To initialize this process, we assume a standard normal prior distribution for $\theta_{i0}$. We also assume $|\rho| < 1$ to ensure stationarity. Thus, the model balances short-term changes in health status with information from long-term health trajectories. Finally, to make the model identifiable in our setting with a binary observation process, we fix $\sigma_e^2 = 1$.

Under the latent variable representation defined by (5), the indication time for each unit $i$ can then be expressed as a function of the observation process

$$T_i = \inf\{t \in [0, K] : \theta_{it} + X_i \beta + \nu_t > 0, \}$$

where $\beta \in \mathbb{R}^p$ is a vector of regression coefficients and $\nu_t \sim N(0, 1)$. One of the main advantages of this representation is that it can flexibly accommodate both fixed and time-varying covariates. The proposed modeling approach can also be extended to settings with non-linear covariate effects by replacing the linear effects with nonlinear contributions (Denison et al. 2002).

We consider a separate model for the assignment mechanism, which determines assignment to treatment versus control upon indication. Here, our model formulation is based on the assumption that, in many settings, variation in treatment practices may be due to clinician and/or institutional preferences rather than differences in patient characteristics (Slaughter et al. 2017). In particular, we assume that each unit is assigned to treatment with a probability that depends on their indication time. Conditional on $T_i$, and on exogenous covariates $D_i$, the assignment mechanism for our specific model can be expressed as

$$Z_i \sim \text{Bernoulli}(\pi_{iT}), \quad \text{logit}(\pi_{iT}) = \delta_0 + f(T_i, D_i, \delta_1),$$

where $\delta_0$ is a parameter for the baseline probability of treatment and $f(\cdot|\delta_1)$ is a parameterized, possibly non-linear, deterministic transformation of the indication time (e.g., calendar year or month). This model can be easily extended to accommodate other artifacts of the study that are believed to influence the probability of treatment assignment. For example, in Sect. 5, we let $\text{logit}(\pi_{iT}) = \delta_0 + \delta_1 D_i$ where $D_i$ is calendar time, and $\delta_1$ is restricted to be negative so that the probability of receiving the treatment is monotonically decreasing over time.

4.2 Inferential procedure

Our procedure uses the Kalman Filter (Carlin and Polson 1992) to marginalize out the latent state parameters $\theta_{1:T}$ for more efficient inferences. The full likelihood of the parameters $\Theta = (\rho, \beta, \delta_0, \delta_1)$ can be written as
where

\[ p(M_i = 0|\Theta, X) = 1 - \sum_{t=0}^{K} p(T_{i}^{mis} = t|\Theta, X)p(Z_i = 1|T_{i}^{mis} = t, \Theta). \]

Letting \( Z \) and \( M \) be the vectors of the \( Z_i \) and \( M_i \) across the \( N \) patients, the associated posterior density is

\[ p(\Theta|T^{obs}, Z, X, M) \propto p(\Theta) \prod_{i=1}^{N} p(T_{i}^{obs}|X, \Theta)p(Z_i|T_{i}^{obs}, \Theta)p(M_i = 0|\Theta, X). \]

Our modeling framework permits a flexible choice of prior distributions. For example, we can assume a prior distribution that factors into independent densities, as

\[ p(\Theta) = p(\rho)p(\beta)p(\delta_1)p(\delta_0), \]

which allows for many distributional choices for each prior component.

Posterior sampling is then straightforward to implement using standard software for implementing MCMC methods such as JAGS (Plummer et al. 2003). The resulting posterior samples allow us to measure the missing times of indication and assignment to control according to (6). Given the inferred indication times, our framework implies that \( Z_i = 0 \) for units with \( T_i \in [0, K] \). Recall that treatment assignment is undefined for those whose indication time was censored by the end of the study (i.e. units with \( T_i > K \)). Finally, we can calculate the observed outcomes for units inferred to belong to the group of true controls by computing the difference between the observed event times and the inferred indication times for those units.

### 4.3 Bayesian analysis with inferred indication times

One of the benefits of the proposed approach is that it allows us to make inferences about the missing indication times that are free from the outcome analysis. Our approach allows for flexible specification of the causal estimands of interest and also allows researchers to choose any mode of inference for analysis of the outcomes that they see appropriate. For example, one might use the posterior mode of inferred times of indication for each untreated unit calculated over a large number of MCMC samples as the point estimate for that unit’s indication time. This can then be viewed as a single imputation of the missing values, and conditional on these estimates one could estimate the treatment effects by simple comparisons of means of the time-to-event outcomes (e.g., using a Neymanian or Fisherman mode of inference). Alternatively, our framework also allows for more sophisticated analysis of outcomes. For instance, we could first obtain a large number of posterior samples for the missing times of indication across all untreated units and use these samples to form unit-level empirical posterior distributions of the missing indication times. Then, one might iteratively sample values from these distributions and calculate the corresponding observed potential outcomes in each iteration. Contrasts between treatment and control...
It is important to note that our proposed approach ignores the time-to-event outcome data when making inferences about the unobserved indication times and treatment assignments. Instead, following Rubin’s (2010) template for the design of observational studies in comparative effectiveness research, we separate the inferential procedure of estimating the indication times from the eventual outcome analysis. In doing so, we aim to preserve the integrity of the underlying hypothetical randomized experiment, whereby the study design takes place prior to collecting any outcome data (Rubin 2007). However, for other applications the model could be altered by extending the indication time likelihood with factors involving the time-to-event outcomes and determining the posterior density in one step.

5 An application to the study of inappropriate prescribing practices for treatment of pulmonary hypertension

To illustrate our proposed methods, we analyzed data from a recent study using electronic medical records from the VA health system evaluating the impacts of inappropriate prescribing practices for the treatment of pulmonary hypertension (Kim et al. 2018). Pulmonary hypertension is a condition of high blood pressure that affects arteries in the lungs and heart. One common treatment for PH is a class of medications called PDE5Is, which act on enzymes causing blood vessels to relax in order to lower blood pressure. While PDE5Is have been shown to be effective for treating some rare forms of PH, recent studies have identified the use of these drugs as wasteful, ineffective, and potentially harmful for treating patients with more common types of PH caused by left heart disease (WHO Group 2) or hypoxemic lung disease (WHO Group 3) (Freiman et al. 2015). Despite its contraindication, a study of veterans diagnosed with these types of PH over the years of 2005 to 2012 identified over 2,000 prescriptions for PDE5Is that were inappropriately administered to patients in the VA health system. To understand the impact of these inappropriate treatment practices on patient outcomes, it is of interest to measure the causal effects of prescribing PDE5Is to patients with PH Groups 2 and 3 on the time lag between the application of the intervention and the occurrence of a clinical event of interest (e.g., time from treatment to death). The primary question that we sought to address through this analysis concerned the extent to which treatment with PDE5Is for patients diagnosed with PH Groups 2 and 3 impacted the mean survival time compared to patients who had similar indications for treatment but were not prescribed PDE5Is during the study period.

5.1 Data

Our data set contains demographic and laboratory measurements as well as records of the utilization of medications, inpatient and outpatient services for over 350,000 veterans who were diagnosed with PH Groups 2 and 3 and received prescription medications from the VA between the years of 2005 to 2015. For all patients, baseline health-related measurements were collected at the time of first PH diagnosis based on ICD-9 diagnosis codes. Subsequent health-related measurements were collected at intermittent observation times corresponding to patient-provider clinical interactions (e.g., an inpatient or outpatient visit). The exact number of measurements recorded and the time elapsed between
Table 2  Bias, RMSE, and coverage rates (CR) of 95% credible intervals for each parameter

| n   | $\pi_X$ | $q_1$ | $q_2$ | $p_1$ | $p_2$ | $p_3$ |
|-----|---------|-------|-------|-------|-------|-------|
|     | Bias RMSE CR | Bias RMSE CR | Bias RMSE CR | Bias RMSE CR | Bias RMSE CR | Bias RMSE CR |
| 100 | 0.1 | 0.03 0.14 0.99 | 0.3 0.12 0.98 | 0.6 −0.02 0.11 0.99 | 0.9 −0.04 0.10 1.00 | −0.01 0.12 0.99 | 0.02 0.13 0.96 | 0.05 0.15 0.96 |
|     | 0.3 | 0.01 0.12 0.98 | 0.01 0.08 0.98 | 0.03 0.11 0.98 | 0.05 0.10 0.99 | 0.07 0.15 0.90 | 0.03 0.13 0.98 |
|     | 0.6 | −0.02 0.11 0.99 | 0.00 0.09 0.98 | 0.03 0.11 0.98 | 0.06 0.14 0.91 | 0.05 0.15 0.96 |
|     | 0.9 | −0.04 0.10 1.00 | 0.01 0.13 0.99 | 0.05 0.10 0.99 | 0.07 0.15 0.90 | 0.03 0.13 0.98 |
| 200 | 0.1 | 0.02 0.14 0.99 | 0.01 0.08 0.99 | 0.01 0.12 0.98 | 0.04 0.12 0.98 | 0.05 0.12 0.95 | 0.06 0.13 0.92 |
|     | 0.3 | −0.02 0.12 0.98 | 0.00 0.07 0.97 | 0.04 0.12 0.98 | 0.05 0.12 0.96 | 0.06 0.12 0.93 | 0.06 0.13 0.93 |
|     | 0.6 | −0.03 0.11 0.97 | −0.01 0.08 0.97 | 0.05 0.12 0.97 | 0.06 0.12 0.93 | 0.06 0.13 0.93 |
|     | 0.9 | −0.04 0.10 0.99 | −0.01 0.09 0.98 | 0.05 0.10 0.99 | 0.06 0.11 0.93 | 0.04 0.12 0.96 |
| 500 | 0.1 | −0.02 0.13 0.99 | 0.00 0.07 0.97 | 0.04 0.12 0.97 | 0.04 0.11 0.96 | 0.07 0.11 0.92 |
|     | 0.3 | −0.03 0.11 0.95 | −0.02 0.07 0.95 | 0.05 0.13 0.95 | 0.05 0.11 0.94 | 0.06 0.11 0.90 |
|     | 0.6 | −0.03 0.10 0.95 | −0.01 0.07 0.94 | 0.05 0.12 0.94 | 0.04 0.09 0.94 | 0.05 0.10 0.92 |
|     | 0.9 | −0.04 0.11 0.98 | −0.03 0.08 0.95 | 0.06 0.12 0.96 | 0.05 0.09 0.95 | 0.03 0.10 0.97 |
| 1000| 0.1 | −0.03 0.12 0.96 | −0.02 0.07 0.94 | 0.06 0.13 0.96 | 0.05 0.11 0.95 | 0.06 0.10 0.91 |
|     | 0.3 | −0.03 0.09 0.90 | −0.01 0.06 0.91 | 0.04 0.10 0.95 | 0.04 0.09 0.95 | 0.04 0.08 0.94 |
|     | 0.6 | −0.02 0.09 0.90 | −0.01 0.06 0.93 | 0.03 0.10 0.93 | 0.03 0.07 0.95 | 0.03 0.08 0.95 |
|     | 0.9 | −0.03 0.10 0.96 | −0.02 0.07 0.96 | 0.04 0.11 0.96 | 0.04 0.06 0.99 | 0.02 0.09 0.96 |
subsequent measurements varied by patient. Observations with implausible values for lab measurements or demographic variables were excluded prior to analysis.

For the present analysis, we considered patients who at the time of PH diagnosis (Group 2 or 3) were between 65 to 95 years of age, were eligible for Medicare benefits, and who had not received prescriptions for a PDE5I medication prior to PH diagnosis. Because female patients comprised less than 3% of the patient population who met these initial eligibility criteria, we further limited our analysis to only male patients. For data integrity purposes, we also excluded any patients who were receiving Medicare part C at the time of diagnosis as well as patients who did not fill any prescriptions within the VA health system during the one year prior to their diagnosis, since these patients may have received PH related care from private providers that we could not track. Finally, we excluded any patients who received a prescription for nitrates in the one-year period following PH diagnosis because nitrates are contraindicated for treatment with PDE5Is. Among patients who met all initial eligibility criteria, we defined the “treatment group” as the set of patients who filled at least one prescription during the study period for a PDE5I medication primarily indicated for the treatment of PH. In particular, we excluded medications with secondary or off-label indications for the treatment of PH such as sildenafil, tadalafil, and vardenafil. Because we observed the prescription dispense dates rather than the dates when the medications were first prescribed, we further excluded from our analyses any treated patient whose first PDE5I prescription was dispensed more than 60 days after a hospital visit. After applying all exclusion restrictions, the remaining sample was comprised of 531 patients who received treatment for PH with a PDE5I medication within a one-year period following their diagnosis date and 167,701 potential controls who did not receive a PDE5I during that period.

Our outcome of interest is survival time in the period following indication for treatment, which we observed for treated units and must be inferred for units in the control group. We consider the date of PH diagnosis as the time of “earliest eligibility” for indication. For each potential control patient, we base our inferences on clinical data observed at intermittent intervals from one year prior to PH diagnosis until death or the end of follow-up in December 2016, whichever occurred first. In both treatment and control groups, patients who survived beyond December 2016 had their survival times censored.

5.2 Constructing comparison groups and time-varying covariates

To make the assumptions of our proposed framework described in Sect. 3 more plausible in this setting, we selected a set of potential control units who appeared similar to the treatment group at baseline based on health-related measurements. Our initial feature set included over 150 baseline covariates capturing various demographic variables, health measurements, and health care utilization records collected in the 1 year prior to PH diagnosis (excluding the date of diagnosis). To reduce the dimensionality of the covariate space, we first performed random forest analysis (Breiman 2001) implemented using the randomForest package in R (Breiman and Cutler 2003) to evaluate the relative importance of each of the observed covariates for predicting indication times among the set of 531 treated units. Using this procedure, we selected a set of 17 baseline covariates identified as most influential for predicting the time of assignment to treatment based on an analysis of variable importance. Among these covariates were several related to the physical characteristics of the patient (e.g., age, height and weight). Other important covariates included the number of comorbidities present at the time of PH diagnosis as well as
variables capturing changes in health care utilization in the 30 days prior to PH diagnosis such as recent organ failure events, recent inpatient visits, or receipt of recent incidental medical procedures.

Potential controls were selected using 1:1 nearest-neighbor matching with replacement, whereby each treated unit is matched to its closest control unit based on the Mahalanobis distance calculated over baseline covariates (Rubin 1973). The reason we prefer matching on the Mahalanobis distance metric, which is essentially a standardized version of the Euclidean distance that weights covariates according to their relative influence on the total variation across the sample, is that it is scale invariant and explicitly adjusts for the correlation between covariates. Compared to alternative strategies such as propensity score matching, pair matching on the Mahalanobis distance between covariates also allows for stricter control over the balance between individual covariates.

This produced a final sample of 531 treated units and 531 matched control units who were similar to the treatment group at their times of PH diagnosis but did not receive a PDE5I at any point during the observation period. Within this sample, survival times were censored at the end of follow up for 140 treated patients (26.4%) and 208 matched controls (39.6%), with the remaining patients dying before the end of the 2016 calendar year. Diagnostics performed after matching confirmed that the covariate distributions were adequately balanced between the treatment group and the matched potential control group. Table 3 shows descriptive statistics on baseline variables for the final matched samples. After matching, we proceeded under the assumption that overall health status was unconfounded given baseline covariates such that patients with similar covariates at baseline were expected to have similar health trajectories and therefore, similar indication times over the course of study.

In addition to baseline covariates, we also included in our analyses a number of time-varying covariates collected at intermittent observation times throughout the study period. Specifically, we considered six time-varying, clinically meaningful covariates that indicate changes in patients’ disease severity and health status over time, including an indicator for whether the patient was most recently observed in an outpatient versus inpatient setting, an indicator for the presence of new comorbidities, and separate indicators for recent hospitalization, organ failure events, cardiac events or pulmonary procedures recorded during the 30 days prior to each visit. We also included as a single time-varying covariate the Mahalanobis distance calculated between each patient’s laboratory measurements (e.g., heart rate and blood pressure) at baseline and values of the same laboratory variables collected at each follow-up visit. This allowed us to greatly reduce the dimension of the covariate space and operationalize the laboratory variables in terms of gain scores.

5.3 Results

Our main analysis involved estimating the impact of prescribing versus withholding contraindicated PDE5I’s on survival time for patients whose indication times occurred within one year of PH diagnosis. Using the approach described in Sect. 3, we fit the model defined by Eqs. (5)–(7) using MCMC posterior simulation to infer the unobserved indication times within the matched control sample. Details of the posterior sampling scheme can be found in Appendix A.

Our health process model component assumed that changes could occur daily, so that \( K = 365 \). Relatively non-informative priors were assumed for the parameters in the model. These priors were as follows:
where $N(-\infty, 0]$ denotes the truncated normal distribution on the interval $(-\infty, 0]$. Here, we constrained $\delta_1$, the effect of calendar time on the log-odds of PDE5I initiation, to be negative for this application since the probability of receiving PDE5Is was believed to be monotonically decreasing over the course of the study. This is because PDE5Is were believed to be the standard, first-line therapy for treatment of PH at the beginning of the study with use steadily decreasing as knowledge of its contraindication for some PH patients spread throughout the medical community.

We ran the MCMC sampler with four parallel chains each run for 20,000 iterations, where the first 5,000 draws of each chain were discarded as a burn-in period. With the resulting 60,000 samples, we calculated the posterior means and 95% credible intervals (CI) for all model parameters. In all cases, the MCMC simulated model parameters and quantities of interest raised no concerns using standard diagnostics including those by Geweke (1992) and Gelman and Rubin (1992). As an additional sensitivity check, we evaluated the performance of the proposed model under different choices of hyperparameters using the deviance information criterion (DIC; Spiegelhalter et al. 2002) and found the results to be generally unaffected.

After convergence was achieved, we continued our MCMC sampling to generate 1,000 posterior draws of the unobserved indication times in the matched control sample. These samples, combined with the observed indication times in the treatment group, were then used to conduct survival analysis. For each posterior sample, we first identified the subset of patients whose indication times (either observed or inferred) were within the specified study period of $K = 365$ days following PH diagnosis. Patients with inferred indication times outside this range were classified as ineligible controls and excluded from the subsequent analysis. Among the remaining set of treated and inferred control patients, we calculated the time-to-event outcome for each patient as the time from indication to death. For patients who survived until the end of the observation period in 2016, survival times were censored at the end of follow-up. This procedure produced, for each posterior sample, a “complete” data set that could then be used to compare survival outcomes between treatment and control groups. For each posterior sample, we then fit a Bayesian proportional hazards model that regressed the hazard of death on an indicator for treatment assignment (Zhou and Hanson 2018; Zhou et al. 2020). Finally, we pooled the results from each of these 1,000 posterior samples to generate estimates, including credible intervals (CIs), for the effects of interest.

Figure 1 shows the estimated survival curves and estimated difference in survival probabilities between treatment (PDE5I) and control groups over the ten-year period following indication for treatment. The posterior mean for the number of “true controls” inferred from the matched sample of potential controls was 457 (95% CI: 118–525). This indicates that many of the potential control units identified by matching on baseline covariates were, in fact, ineligible to receive treatment during the one-year period following their PH diagnosis. For these patients, lack of treatment can therefore be interpreted as a lack of indication for treatment. On the other hand, potential controls with inferred indication times occurring at or before each specified time point are regarded as “true controls” for whom,
upon indication for treatment, PDE5Is were actively withheld, possibly in favor of alternative medication or treatment strategy. Given the inferred times of indication, these patients provide a credible comparison group with which we can compare survival outcomes. In particular, our findings suggest that among patients with PH (Groups 2 and 3) who are indicated for treatment at any point in the one year following their diagnosis, treatment with PDE5Is leads to a significant increase in the mean hazard. The posterior mean hazard ratio for receipt of PDE5Is upon indication among eligible PH patients was 1.58 (95% CI: 1.32–1.94).

Table 4 shows posterior estimates for all model parameters including the correlation between patients’ latent health states, $\rho$, the parameters governing the probability of assignment to treatment upon indication, $\delta_0$ and $\delta_1$, and the baseline and time-varying covariate effects, $\beta_1$ and $\beta_2$. We present effect estimates only for the set of baseline and time-varying covariates identified as having 95% credible intervals not containing 0. Overall, we find a posterior mean autocorrelation $\rho$ near 1 for latent health measurements over time. This suggests that the latent health process is nearly a Gaussian random walk without any notable regression towards the mean. Based on the posterior estimates for parameters $\delta_0$ and $\delta_1$, our results suggest a baseline probability of treatment upon indication of approximately 0.86 for patients diagnosed with PH in 2006, with the probability of treatment slowly decreasing over time.

A key feature of our model is that it also allows us to directly evaluate which of the baseline and time-varying covariates carry more or less information about patients’ indication times. In general, our findings suggest that patients with PH Group 2 were more likely to have indications for treatment shortly after PH diagnosis than patients with PH Group 3. Results also indicate that the number of comorbidities present at PH diagnosis is strongly associated with earlier indication times. Further, we found that patients who regularly receive care in an inpatient setting generally have earlier indication times than patients who receive routine care in an outpatient setting. Among the other baseline covariates included in our analysis, occurrence of one or more incidental medical procedures within 30 days prior to PH diagnosis was also positively associated with indication for treatment during the study period. These results may offer insights for clinicians about best practices for health management of PH patients, and may also be used to guide modeling decisions in other applications.

5.4 Comparison with inferences based on risk set matching

As discussed in Sect. 2, risk set matching (RSM) is a strategy for making causal inferences in longitudinal observational studies that can provide an alternative measure for understanding the impacts of unknown indication times. RSM is, fundamentally, a data processing strategy that aims to identify matched sets of treated and not-yet-treated patients whose outcomes can be compared in order to estimate treatment impacts. Unlike our proposed approach, which aims to assess the effects of initiating versus withholding treatment at the time of treatment indication on a time-to-event outcome, RSM focuses on the effects of initiating versus delaying treatment at a given time point on an outcome observed at a fixed end-point. As a result, effect estimates produced using RSM are not directly comparable to those based on our proposed inferential approach. Nevertheless, RSM can serve as a useful reference for contextualizing our empirical results. Below, we present the results from applying RSM to our study of inappropriate prescribing practices for the treatment of PH.
and discuss how these results can be interpreted relative to those based on our proposed approach.

For each treated patient with observed indication time $t$, we identified the not-yet-treated patient that was most similar in terms of their probability of receiving treatment based on covariates observed up to time $t$. With the resulting risk-set matched sample, we then calculated the (possibly right-censored) time-to-event outcome for each patient as the time from indication to death based on the observed times of indication in the treatment group. We then fit a Bayesian proportional hazards model to estimate the hazard of death over the study period for patients in each risk set.

Figure 2 shows the estimated survival curves and estimated difference in survival probabilities between the treatment (PDE5I) and risk-set matched control groups over the ten-year period following indication for treatment. Similar to previous findings, these results suggest a significant decline in both long-term and short-term survival outcomes associated with treatment. The posterior mean hazard ratio for receipt of PDE5Is upon indication in the risk-set matched sample was 1.39 (95% CI: 1.20–1.62).

As seen in the comparison between Figs. 1 and 2, we find that estimates obtained using RSM are systematically smaller in terms of both magnitude and posterior uncertainty than

| Variable                                | Potential Control | Treatment  |
|-----------------------------------------|-------------------|------------|
| PH Type                                 |                   |            |
| Group 2                                 | 477 (89.8%)       | 477 (89.8%)|
| Group 3                                 | 54 (10.2%)        | 54 (10.2%) |
| Recently Hospitalization                 |                   |            |
| Yes                                     | 73 (13.7%)        | 74 (13.9%) |
| No                                      | 458 (86.3%)       | 457 (86.1%)|
| Recent Incidental Medical Procedure     |                   |            |
| Yes                                     | 424 (79.8%)       | 425 (80.0%)|
| No                                      | 107 (20.2%)       | 106 (20.0%)|
| Recent ER Visit                         |                   |            |
| Yes                                     | 309 (58.2%)       | 309 (58.2%)|
| No                                      | 222 (41.8%)       | 222 (41.8%)|
| Age                                     | 74.7 (6.6)        | 74.8 (6.9) |
| Cardiac Events                          | 1.5 (1.4)         | 1.5 (1.4)  |
| Diastolic Blood Pressure                | 71.6 (10.0)       | 71.3 (11.7)|
| Height                                  | 69.3 (2.4)        | 69.3 (2.6) |
| Inpatient Days                          | 4.0 (7.6)         | 4.9 (11.1) |
| Number of Comorbidities                 | 0.4 (0.8)         | 0.4 (0.7)  |
| Number of Medications                   | 12.4 (5.1)        | 12.4 (5.4) |
| Organ Failure Events                    | 1.2 (1.1)         | 1.3 (1.2)  |
| Outpatient Days                         | 32.6 (20.0)       | 35.2 (22.7)|
| Pulmonary Events                         | 0.7 (0.9)         | 0.8 (1.0)  |
| Resting Heart Rate                      | 76.2 (13.3)       | 76.0 (14.4)|
| Systolic Blood Pressure                 | 130.2 (16.9)      | 129.5 (18.6)|
| Weight                                  | 202.4 (38.3)      | 201.2 (40.1)|

Values shown represent frequencies (%) for categorical variables and means (SD) for continuous variables.
results based on the proposed approach. However, because estimates based on the risk-set matched sample reflect the average effect of treatment upon indication for the treated rather than the average treatment effect across the entire patient population, these differences are not surprising. In particular, estimates calculated using RSM reflect the effects of delaying treatment with PDE5Is for the subset of patients that have indications for treatment during the study period. This means that, for each time \( t \) in the study period, patients treated with PDE5Is at that time are compared to a matched control group consisting of patients who had not yet received treatment at time \( t \) but may have received treatment at a later point in the study. For example, in the present application, we found that 277 (42.7%) of the 531 controls identified using RSM were patients who received treatment at some point in the study period. The main differences between the two approaches therefore lies in the interpretation of the parameters. Where the model proposed in Sect. 4 yields an estimate that captures the impact of assignment to treatment versus control at time \( t \) on time to death for all individuals in the population who had indications for treatment at that time, inferences based on RSM yield an estimate of the time to death for individuals who were indicated versus not yet indicated for treatment at time \( t \), given the confounders measured up to that time.

6 Discussion

In this paper, we propose a novel conceptualization of longitudinal observational studies with treatment by indication and provide a template for the design and analysis of such studies in a manner that approximates a randomized experiment with a binary treatment. This conceptualization reformulates the problem of evaluating a time-varying treatment in the presence of time-varying confounders as one of evaluating a binary treatment administered at a fixed time point, where the time of assignment is a partially observed pre-treatment confounder. Our hope is that this simplified representation of a traditionally complex data structure will allow for more straightforward analyses of health data in the digital age (e.g., data from electronic medical records).

The merit of the proposed model is that it allows for model-based assessments of the times of treatment indication based on disease progression that can accommodate time-varying covariates measured at intermittent observation times (i.e., when there is missing data and/or substantial variation in the timing of patient-provider visits). Our approach allows for systematic evaluation of the underlying health factors that may be most influential in determining a patient’s need for treatment. Inferences based on this modeling strategy may be useful to address a number of important questions in health services research—for example, what preventative health strategies might be most effective for delaying the onset of indications for treatment? The model also allows us to obtain conditional probabilities for a subject having indications for treatment at different points over time, which could be used to inform preventative treatment strategies. We note that the proposed model assumes that subjects’ overall health fluctuates under natural conditions.

As described in Sect. 4, the proposed model for time of indication for treatment can accommodate both fixed and time-varying covariates, which can be useful in explaining differences in the aspects of health associated with subject-specific characteristics and/or conditions that vary between hospital visit within a subject. Alternatively, covariate data could be excluded entirely from the model to make inferences about indication times that
are viewed as fully stochastic. Another possibility for modeling variability in the parameters involves the inclusion of provider-level, or geographic-specific random effects. However, this complicates the evaluation of the likelihood, rendering model fitting more challenging.

Finally, we note that our application and inferential results should be regarded as an illustrative example of the proposed methods rather than an attempt to provide definitive answers about the causal effects of inappropriate prescribing practices on health-related outcomes for PH patients. In particular, for researchers interested in drawing causal inferences from observational studies with time-varying exposures, which often require a challenging design process that can obfuscate resulting inferences, the conceptual framework we present offers an alternative formulation of the underlying problem that is both intuitive and relatively straightforward to implement. The proposed approach can also be flexibly extended to accommodate a range of data structures, for instance in studies with multivariate outcomes.

Appendix A Posterior sampling

General sampling scheme

We employ a Gibbs sampler to draw posterior samples from the model described in Sect. 3. In each iteration of the Gibbs sampler, we draw the missing indication times $T$ for units with $M = 1$ from the conditional posterior predictive distribution of $T$ given covariates $X$ and the current draw of the parameter $\theta$. The completed indication times can then be used to classify untreated patients into distinct groups of true controls and ineligible controls based on eligibility $S$, where the true control group consists of patients with $M = 1$ and $S = 1$. For the true controls, we can then calculate values for the potential outcomes $Y(0)$ given the generated values of $T$. These values are then regarded as the observed potential outcomes, $Y^{obs} = (1 - Z)Y(0) + ZY(1)$. We can then update the parameters $\theta_1$, $\theta_2$ and $\theta_3$ by drawing from the conditional posterior distribution with density function $p(\theta_1, \theta_2, \theta_3 | Y^{obs}, T, M, X)$.

Posterior inference on the causal effects of interest can be obtained by computing the values of the constructed estimator within each MCMC iteration and summarizing their

![Fig. 1 Estimated survival curves for treatment and inferred control groups (left) and estimated difference in survival probabilities (right) with 95% credible interval bands](image-url)
distribution across the posterior sample. Thus, in each iteration, we can construct a dataset consisting of the observed indication times, the simulated indication times, and all observed potential outcomes, and use these completed data to calculate an estimate of the treatment effect. Alternatively, we could specify a joint distribution for the potential outcomes $Y = (Y(0), Y(1))$ that we could then use to impute the missing potential outcomes $Y_{\text{mis}}$ in each iteration by drawing from the conditional distribution with density function $p(Y_{\text{mis}}|Y_{\text{obs}}, T, X, \theta)$. Repeating this process over many such simulated datasets produces the approximate posterior distribution for all causal effects of interest. In the same way, posterior samples of $\theta$ can provide posterior estimates of the parameters that characterize the data-generating process; this is described in greater detail in Sect. 4.

**Full conditionals**

For the model described in Sect. 4.1, we employ the Gibbs sampler as the posterior sampling strategy. Using a data augmentation approach, we also let $\Psi_{it} = \theta_{it} + X_{it}\beta + v_{it}$ where $v_{it} \sim N(0, 1)$ such that the indication times $T_i$ can be represented as $T_i = \inf\{t \in [0, K]: \Psi_{it} > 0\}$. Note that the latent variables $\Psi_{it}$ are conditionally independent across units $i$ and over $t$ given $\theta_{it}$ and $\beta$.

To begin, we set $j = 0$ and draw initial values for the parameters $\Theta = (\rho, \beta, \delta_0, \delta_1)$ and the latent variables $\theta_{1:K},\Psi_{1:K}$ for all $i = 1, \ldots, n$. The latent variables $\Psi_{1:K}$ are then used to determine the initial values for the missing indication times $T_i$. For each iteration, we then proceed as follows:

(a) Draw the latent variables $\Psi_{it}$ from the full conditional distribution given below, which is either an unrestricted normal density (when $T > t$) or a truncate normal density on the interval $(-\infty, 0]$ when $T < t$ or $(0, \infty)$ when $T = t$.

$$p(\Psi_{1:K}|\cdot) = \prod_{i=1}^{N} \prod_{t=1}^{K} 1(\Psi_{it} > 0)^{T_i < t} 1(\Psi_{it} \leq 0)^{T_i \geq t} \phi(\Psi_{it} - \theta_{it} + X_{it}\beta)$$

| Parameter | Mean | SD  | 95% CI   |
|-----------|------|-----|----------|
| $\beta_1$ (Baseline covariates) | | | |
| PH Group 2 | $-4.55$ | $1.12$ | $(-6.33, -2.27)$ |
| No. comorbidities | $-0.76$ | $0.23$ | $(-1.22, -0.31)$ |
| Recently hospitalized | $1.06$ | $0.37$ | $(0.33, 1.75)$ |
| Recent procedures | $-1.02$ | $0.39$ | $(-1.72, -0.25)$ |
| $\beta_2$ (Time-varying covariates) | | | |
| Outpatient visit | $1.77$ | $0.44$ | $(0.79, 2.41)$ |
| No. new comorbidities | $0.32$ | $0.14$ | $(0.06, 0.58)$ |
| New hospitalizations | $0.54$ | $0.24$ | $(0.08, 1.02)$ |
| Change in labs | $0.13$ | $0.05$ | $(0.04, 0.21)$ |
| $\rho$ | $0.98$ | $0.01$ | $(0.96, 0.98)$ |
| $\delta_0$ | $1.88$ | $0.16$ | $(1.58, 2.19)$ |
| $\delta_1$ | $0.04$ | $0.02$ | $(0.00, 0.08)$ |
where \( \phi(\cdot) \) denotes the probability density function of the standard normal distribution.

(b) Since Eqs. (5) and (6) define a linear state space model, we sample \( \theta_{i,1:K} \) using the Kalman filter. The forward conditional is given by:

\[
p(\theta_{1:K} | \cdot) = \prod_{i=1}^{N} \prod_{t=1}^{K} \phi (\theta_{i,t} - \rho \theta_{i,t-1})
\]

where \( \phi(\cdot) \) denotes the probability density function of the standard normal distribution.

(c) Next, draw the probability of assignment to treatment upon indication from the conditional distribution given by:

\[
p(\pi | \cdot) = \prod_{i=1}^{N} \prod_{t=1}^{K} \left( \pi_{it}^{2} (1 - \pi_{it})^{1 - Z_{it}} \right)^{1(T_{it}=1)}
\]

(d) Assuming a general multivariate normal prior for \( \beta \) of the form \( \beta \sim N(\beta_0, \Sigma_0) \), draw \( \beta \) from the multivariate normal distribution \( \beta | \cdot \sim N_p(\beta_1, \Sigma_1) \) where

\[
\beta_1 = \Sigma_1 \left( \Sigma_0^{-1} \beta_0 + \sum_{i=1}^{N} \sum_{t=1}^{K} X_{it} (\Psi_{it} - \theta_{it}) \right).
\]

and

\[
\Sigma_1 = \left( \Sigma_0^{-1} + \sum_{i=1}^{N} \sum_{t=1}^{K} X_{it} X_{it}^T \right)^{-1}
\]

(e) Draw the autocorrelation parameter \( \rho \) from the full conditional distribution given by the truncated normal distribution

\[
\rho | \cdot \sim N_{[-1,1]} \left( \frac{\sum_{i=1}^{N} \sum_{t=1}^{K} \theta_{i,t} \theta_{i,t-1}}{\sum_{i=1}^{N} \sum_{t=1}^{K} \theta_{i,t}^2 \theta_{i,t-1}}, \frac{1}{\sum_{i=1}^{N} \sum_{t=1}^{K} \theta_{i,t-1}^2} \right)
\]

Fig. 2 Estimated survival curves for treatment and risk-set matched control groups (left) and estimated difference in survival probabilities (right) with 95% credible interval bands
Funding This work was supported in part by the United States Department of Veterans Affairs (VA) Health Services Research and Development (IIR 15-115).

Data availability All data and replication materials for the simulation study described in Sect. 3.4 are available at https://github.com/reaganmozer/longbayes. However, all data generated and/or analyzed in Sect. 5 are constructed from confidential patient-level data, which cannot be made available due to restrictions set forth in the data use agreement signed by the authors.

Code availability Pertinent source code related to the findings in Sect. 5 (without the data) is available from the authors upon reasonable request.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Albert, J.H., Chib, S.: Bayesian analysis of binary and polychotomous response data. J. Am. Stat. Assoc. 88(422), 669–679 (1993)
Angrist, J.D., Imbens, G.W., Rubin, D.B.: Identification of causal effects using instrumental variables. J. Am. Stat. Assoc. 91(434), 444–455 (1996)
Basse, G.W., Volfovsky, A., Airola, E.M.: Observational studies with unknown treatment. arXiv preprint arXiv:1601.04083 (2016)
Breiman, L.: Random forests. Mach. Learn. 45(1), 5–32 (2001)
Breiman, L., Cutler, A.: Manual for Setting Up, Using, and Understanding Random Forests. (2003) https://www.stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf
Byar, D.P.: Why data bases should not replace randomized clinical trials. Biometrics 36(2), 337–342 (1980)
Carlin, B.P., Polson, N.G.: Monte Carlo Bayesian methods for discrete regression models and categorical time series. Bayesian Stat. 4, 577–586 (1992)
Danaei, G., Rodríguez, L.A.G., Cantero, O.F., Logan, R., Hernán, M.A.: Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. Stat. Methods Med. Res. 22(1), 70–96 (2013)
Denison, D.G., Holmes, C.C., Mallick, B.K., Smith, A.F.: Bayesian Methods for Nonlinear Classification and Regression, vol. 386. Wiley, New York (2002)
Freiman, M.R., Rose, A.J., Powell, R.W., Miller, D.R., Wiener, R.S.: Patterns of potentially inappropriate prescribing of phosphodiesterase inhibitors in pulmonary hypertension in the va. In: C13. Accounting for costs and resource utilization in respiratory health, pp. A3889–A3889. American Thoracic Society (2015)
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: Bayesian Data Analysis, vol. 2. Chapman & Hall/CRC, Boca Raton (2014)
Gelman, A., Rubin, D.B.: Inference from iterative simulation using multiple sequences. Stat. Sci. 7(4), 457–472 (1992)
Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. 6, 721–741 (1984)
Geweke, J.: Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. Bayesian Stat. 4, 641–649 (1992)
Hernán, M.A., Alonso, A., Logan, R., Grodstein, F., Michels, K.B., Stampfer, M.J., Willett, W.C., Manson, J.E., Robins, J.M.: Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology 19(6), 766 (2008)
Holland, P.W.: Statistics and causal inference. J. Am. Stat. Assoc. 81(396), 945–960 (1986)
Kim, D., Lee, K.M., Freiman, M.R., Powell, W.R., Klings, E.S., Rinne, S.T., Miller, D.R., Rose, A.J., Wiener, R.S.: Phosphodiesterase-5 inhibitor therapy for pulmonary hypertension in the united states. Actual versus recommended use. Ann. Am. Thorac. Soc. 15(6), 693–701 (2018)
Levine, M.N., Julian, J.A.: Registries that show efficacy: good, but not good enough. J. Clin. Oncol. 26(33), 5316–5319 (2008)
Li, Y.P., Propert, K.J., Rosenbaum, P.R.: Balanced risk set matching. J. Am. Stat. Assoc. 96(455), 870–882 (2001)

Mamdani, M., Rochon, P.A., Juurlink, D.N., Kopp, A., Anderson, G.M., Naglie, G., Austin, P.C., Laupacis, A.: Observational study of upper gastrointestinal haemorrhage in elderly patients given selective cyclooxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs. BMJ 325(7365), 624 (2002)

McGettigan, P., Henry, D.: Cardiovascular risk and inhibition of cyclooxygenase a systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2. JAMA 296(13), 1633–1644 (2006)

Plummer, M., et al.: Jags: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing, vol. 124, pp. 1–10. Vienna, Austria (2003)

Poses, R.M., Smith, W.R., McClish, D.K., Anthony, M.: Controlling for confounding by indication for treatment: are administrative data equivalent to clinical data? Medical Care, pp. AS36–AS46 (1995)

Robins, J.M., Hernán, M.A., Brumback, B.: Marginal structural models and causal inference in epidemiology. Epidemiology 11(5), 551 (2000)

Rubin, D.B.: Matching to remove bias in observational studies. Biometrics 29, 159–183 (1973)

Rubin, D.B.: Inference and missing data. Biometrika 63(3), 581–592 (1976)

Rubin, D.B.: Randomization analysis of experimental data: the fisher randomization test comment. J. Am. Stat. Assoc. 75(371), 591–593 (1980)

Rubin, D.B.: William G. Cochran’s contributions to the design, analysis, and evaluation of observational studies. In: Rao, S., Sedransk, J. (eds.) Research Work of William G. Cochran, pp. 37–69. Wiley, New York (1984)

Rubin, D.B.: The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. Stat. Med. 26(1), 20–36 (2007)

Rubin, D.B.: On the limitations of comparative effectiveness research. Stat. Med. 29(19), 1991–1995 (2010)

Slaughter, J.L., Reagan, P.B., Newman, T.B., Klebanoff, M.A.: Comparative effectiveness of nonsteroidal anti-inflammatory drug treatment vs no treatment for patent ductus arteriosus in preterm infants. JAMA Pediatr. 171(3), e164354–e164354 (2017)

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A.: Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B (Stat. Methodol.) 64(4), 583–639 (2002)

van Houwelingen, H., Putter, H.: Dynamic Prediction in Clinical Survival Analysis. CRC Press, Boca Raton (2011)

Watson, D., Spaulding, A.B., Dreyfus, J.: Risk-set matching to assess the impact of hospital-acquired bloodstream infections. Am. J. Epidemiol. 188(2), 461–466 (2019)

Yoshida, K., Solomon, D.H., Kim, S.C.: Active-comparator design and new-user design in observational studies. Nat. Rev. Rheumatol. 11(7), 437–441 (2015)

Zhou, H., Hanson, T.: A unified framework for fitting Bayesian semiparametric models to arbitrarily censored survival data, including spatially referenced data. J. Am. Stat. Assoc. 113(522), 571–581 (2018)

Zhou, H., Hanson, T., Zhang, J.: spbayessurv: Fitting Bayesian spatial survival models using r. J. Stat. Softw. 92(1), 1–33 (2020)

Zhou, Z., Rahme, E., Abrahamowicz, M., Pilote, L.: Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. Am. J. Epidemiol. 162(10), 1016–1023 (2005)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.