Causal Inference with Recurrent Event Outcomes

Estimands, Identification, and Bayesian Inference

Arman Oganisian

Department of Biostatistics
Brown University

ENAR 2025

Analysis of Recurrent Event Outcomes

Biomedical studies often involve outcomes that can recur many times.

Goal: contrast two treatments on the basis of event occurance rate within a defined follow-up window.

Analysis with observational data is challenging:

- Lack of randomization - need to adjust for observed confounding.
- Terminal events - terminal event process stops the recurrent event process.
- Censoring - censoring event process coarsens both recurrent and terminal event process.
- Design - treatment rarely initiated at time zero.

Motivating Example: Hospitalization Risk

Context: Opioids are commonly used to treat chronic back pain (CBP), but could lead to increased hospitalization.

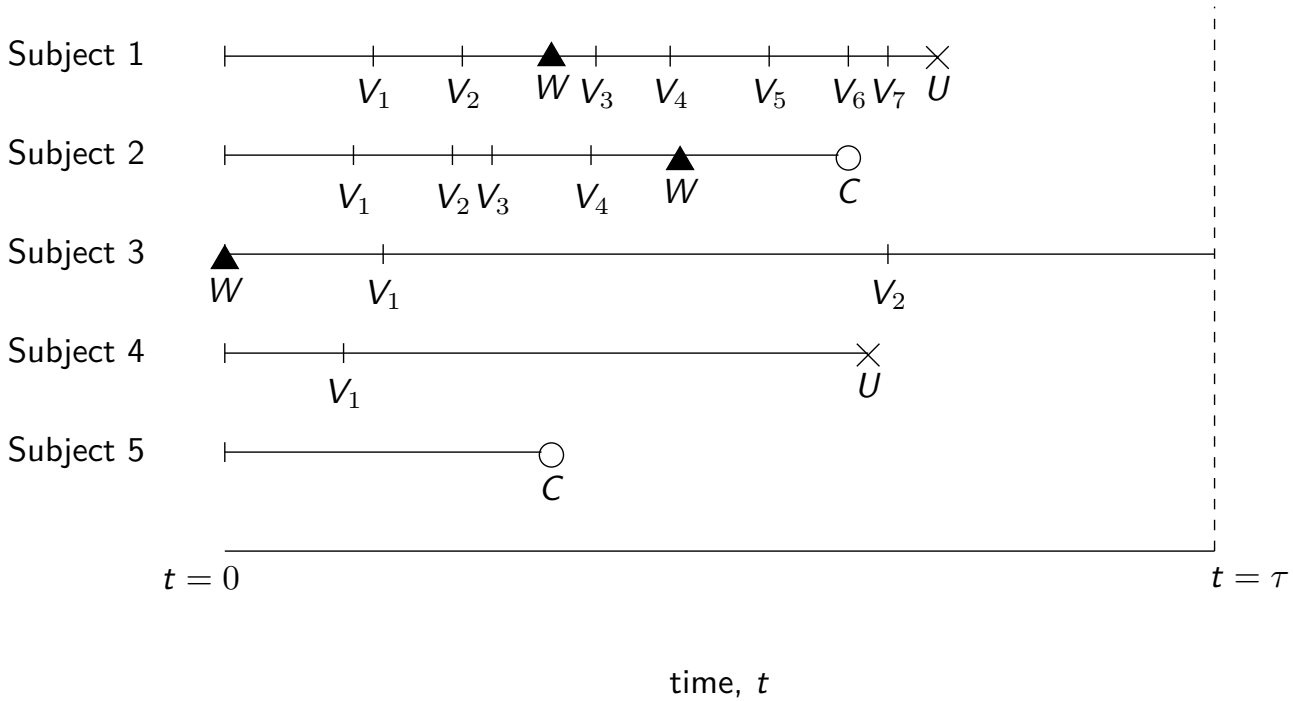
Question: Does opioid therapy increase hospitalization risk among patients with CBP?

- Medicare claims data.
- Target population: patients with CBP who meet eligibility criteria.
- Recurrent event process: hospitalizations.
- Terminal event process: death.
- Censoring event process: loss of medicare, end of data cut.

Talk Outline

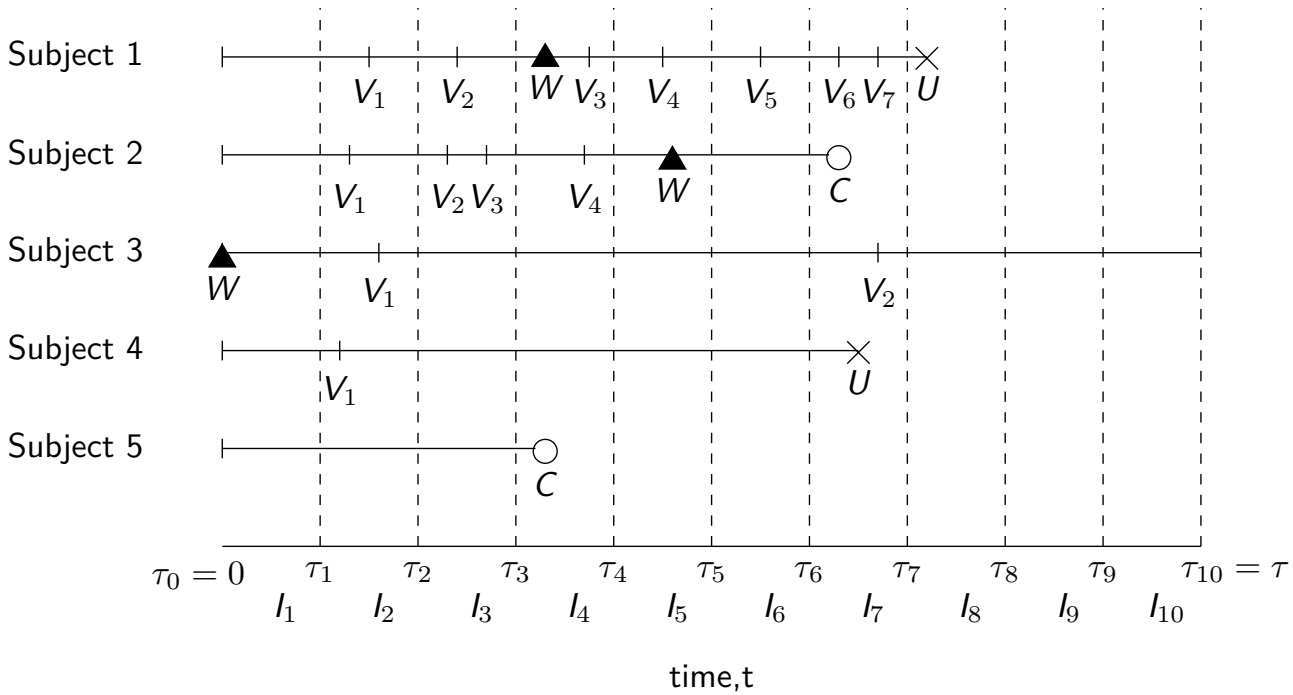
- Data structure.
- Potential outcomes and estimands.
- Bayesian models.

Observed Data Structure



- Time zero: date at which eligibility criteria are met.
- W: time of opioid initiation.
- V_j: time of j^{th} hospitalization.
- U: time of death; C: time of censoring.
- $\tilde{U} = \min(U, C)$.

Discretized Data



- Data discretized into $k = 1, 2, \ldots, K$ intervals.
- A_k: binary, on opioid at interval k.
- Y_k: # events in interval k.
- T_k: binary dead/alive status; C_k: binary censoring status.

Potential Outcomes and Estimands

Treatment Initiation Strategies

To accomodate random initiation times, W, we consider time-varying treatment strategies given by the K-vector

$$a(s) = (\underbrace{0, 0, 0, \ldots, 0}_{\text{entries } 1 \text{ to } s-1}, \underbrace{1, 1, \ldots, 1}_{\text{entries } s \text{ to } K})$$

for $s \in \{1, 2, \ldots, K+1\}$.

There are $K+1$ such strategies.

- E.g. strategy $s = K+1$: never initiate.
- E.g. strategy $s = 1$: initiate at eligibility.
- Can trace out an effect curve across s.

Potential Survival and Recurrent Event Processes

Under strategy $a(s)$,

- $T_k^{a(s)}$: Potential survival status at k had we followed $a(s)$ through k.
- $Y_k^{a(s)}$: Potential event count at k had we followed $a(s)$ through k.

Key constraints:

- Death is an absorbing state: $P(T_k^{a(s)} = 1 \mid T_{k-1}^{a(s)} = 1, \bar{Y}_{k-1}^{a(s)}, -) = 1$
- Death is a terminal event: $P(Y_k^{a(s)} = 0 \mid T_k^{a(s)} = 1, \bar{Y}_{k-1}^{a(s)}, -) = 1$

Censoring is an intervention, but we suppress indexing so $Y_k^{a(s)} = Y_k^{a(s),\bar{c}_k=0}$.

Over/under bars denote history/future:
$\bar{X}_k = (X_1, X_2, \ldots, X_k)$; $\underline{X}_k = (X_k, X_{k+1}, X_{k+2}, \ldots, X_K)$

Marginal Contrasts of Potential Incidence Rate

Difference in expected potential incidence rate had everyone in the target population initiated opioids at time s vs. s'.

$$\Psi(s, s') = E_{P^*}\left[\frac{\sum_{k=1}^{K} Y_k^{a(s)}}{K - \sum_{k=1}^{K} T_k^{a(s)}}\right] - E_{P^*}\left[\frac{\sum_{k=1}^{K} Y_k^{a(s')}}{K - \sum_{k=1}^{K} T_k^{a(s')}}\right]$$

- Mimics usual incidence rates typically used for count outcomes[1].
- Combines information about both event count and survival.
- Can consider other functions, $g(\bar{Y}_K^{a(s)}, \bar{T}_K^{a(s)})$.
- Analysis involves *two* sets of potential outcomes. Inherently difficult to tell a full story with a single estimand!

Requires identification of joint probability mass function P^*.

[1] See Janvin et al. 2023 and Schmidli et al. 2023 for discussion in point-treatment settings

Associational Estimand

Ever-Never Analysis 😱:

$$\log E[Y \mid \tilde{A} = \tilde{a}, L = l, \tilde{U} = \tilde{u}] = \beta_0 + \beta_1 \tilde{a} + l'\beta_2 + \log(\tilde{u})$$

- Y: total event count.
- L: baseline covariates.
- $\tilde{A}$: indcator of ever/never initiated opioid within $[0, \tau]$.
- $\exp(\beta_1)$ is the associational incidence rate ratio:

$$\exp(\beta_1) = \frac{E[Y/\tilde{u} \mid \tilde{A} = \tilde{1}, L = l, \tilde{U} = \tilde{u}]}{E[Y/\tilde{u} \mid \tilde{A} = \tilde{0}, L = l, \tilde{U} = \tilde{u}]}$$

Problems:

- $\tilde{A}$ misattributes person-time and events to treatment.
- Conditions on post-baseline survival, $\tilde{U}$.
- Unclear what implicit assumptions are on censoring.

Alternative Causal Estimand

Survivor-Average Causal Effect (SACE):

$$E\Big[\sum_{k=1}^{K} Y_k^{a(s)} \mid \bar{T}_K^{a(s)} = \bar{T}_K^{a(s')} = 0\Big] - E\Big[\sum_{k=1}^{K} Y_k^{a(s')} \mid \bar{T}_K^{a(s)} = \bar{T}_K^{a(s')} = 0\Big]$$

- A valid causal contrast!
- ...but principal stratum of "always survivors" is not identifiable and may not even be a relevant subgroup.
- Recurrent event process is not "undefined" after death!

Identification Assumptions

Given baseline confounders L, P^* is identified under,

1. Sequential Ignorability:

$$\underline{Y}_k^{a(s)}, \underline{T}_k^{a(s)} \perp C_k, A_k \mid \bar{A}_{k-1}, L, \bar{Y}_{k-1}, C_{k-1} = T_{k-1} = 0$$

2. Sequential Positivity:

$$\underbrace{\lambda_s^A(l, \bar{y}_{s-1})(1 - \lambda_s^C(l, \bar{y}_{s-1}))}_{\text{initiate at } s} \underbrace{\prod_{k=1}^{s-1}(1 - \lambda_k^A(l, \bar{y}_{k-1}))(1 - \lambda_k^C(l, \bar{y}_{k-1}))}_{\text{remain uncensored and un-initiated until } s-1} > 0$$

- λ_k^A and λ_k^C denote hazard of treatment initiation and censoring.
- Can be estimated to guide choice of s.

Identification via G-Formula

Joint pmf, P^*, of $(\bar{Y}_K^{a(s)}, \bar{T}_K^{a(s)})$ is identified as

$$P^*\Big(y(k), t(k)\Big) = \int \lambda_k(\bar{a}_k, \bar{y}_{k-1}, l) \prod_{j=1}^{k-1} f(y_j \mid \bar{a}_j, \bar{y}_{j-1}, l)\big(1 - \lambda_j(\bar{a}_j, \bar{y}_{j-1}, l)\big)\, dF_L(l)$$

1. Discrete-time hazard model:

$$\lambda_k(\bar{a}_k, \bar{y}_{k-1}, l) = P(T_k = 1 \mid T_{k-1} = C_{k-1} = 0, \bar{a}_k, \bar{y}_{k-1}, l)$$

2. Recurrent event model:

$$f(y_k \mid \bar{a}_k, \bar{y}_{k-1}, l) = P(Y_k = y_k \mid T_k = C_k = 0, \bar{a}_k, \bar{y}_{k-1}, l)$$

   with corresponding intensity $\mu_k(\bar{a}_k, \bar{y}_{k-1}, l; \theta)$.

3. Confounder model, $dF_L(l)$.

Expectations $E_{P^*}[g(\bar{Y}_K^{a(s)}, \bar{T}_K^{a(s)})]$ evaluated via Monte Carlo.

Discrete-Time Stochastic Process Models

Discrete-Time Models

Discrete-time hazard of terminal event (death),

$$\lambda_k(\bar{a}_k, \bar{y}_{k-1}, l; \beta) = \text{expit}\left(\beta_{0k} + l'\beta_L + y_{k-1}\beta_Y + \beta_A a_k\right)$$

Intensity function of Recurrent Event.

$$\mu_k(\bar{a}_k, \bar{y}_{k-1}, l; \theta) = \exp\left(\theta_{0k} + l'\theta_L + y_{k-1}\theta_Y + \theta_A a_k\right)$$

λ_k and μ_k characterize the joint evolution of terminal and recurrent event process.

- Parameters $\beta = (\{\beta_{0k}\}_{k=1}^K, \beta_L, \beta_Y, \beta_A)$ and $\theta = (\{\theta_{0k}\}_{k=1}^K, \theta_L, \theta_Y, \theta_A)$.
- Flexible baseline hazard $\{\beta_{0k}\}_{k=1}^K$ and baseline intensity $\{\theta_{0k}\}_{k=1}^K$.
- Larger $K \rightarrow$ more flexability.
- Bayesian inference: priors over $(\beta, \theta) \times$Likelihood $\rightarrow$ posterior over (β, θ).

Estimation is Challenging for Large K

Smoothing is done in ad-hoc, trial-and-error ways. Some examples:

- Young et al (2020): intercept set to be a second-order polynomial function of k "after several bootstrap samples for the construction of confidence intervals failed to converge under the more flexible model."
- Hernán et al. (2000): "We cannot estimate a separate intercepts for each month k. Rather, we need to 'borrow strength' from subjects starting zidovudine in months other than k to estimate $[\beta_{0k}]$. This can be accomplished by assuming that $[\beta_{0k}]$ is constant in windows of, say, 3 months."
- Dodd et al. (2019): "Taking into account the frequency and duration of follow-up information in this analysis with the potential for covariate information to be updated on a daily basis, it seemed sensible to use fortnightly intervals."

Expressing either 1) prior beliefs or 2) need for smoothing.

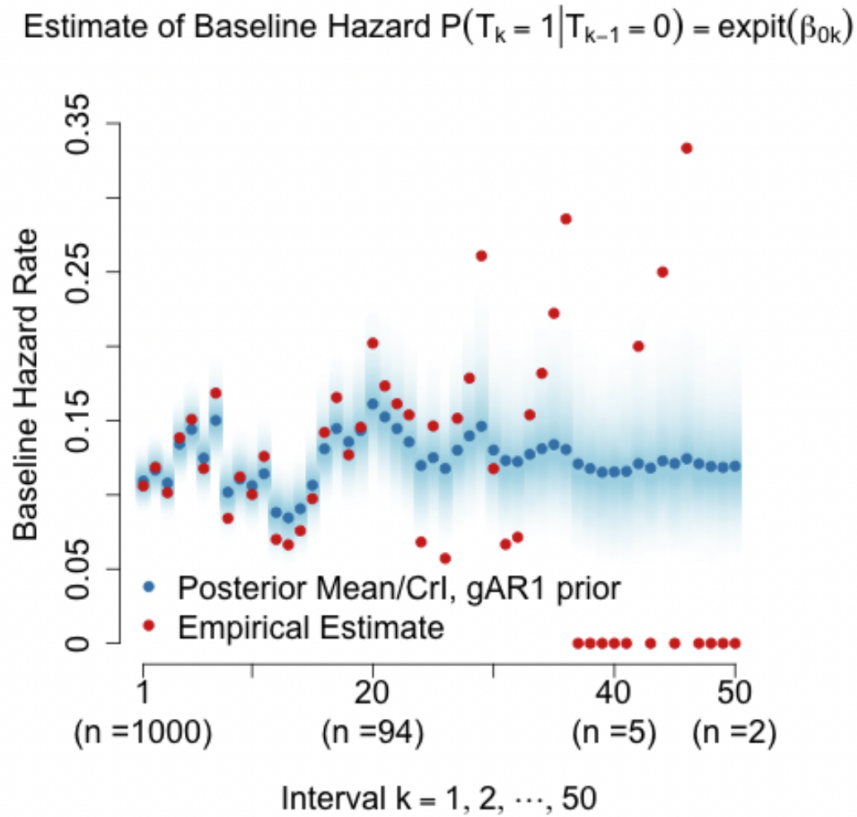
Bayesian Approach: Temporal Smoothing Prior

Use joint process for $\{\beta_{0k}\}_{k=1}^{K}$:

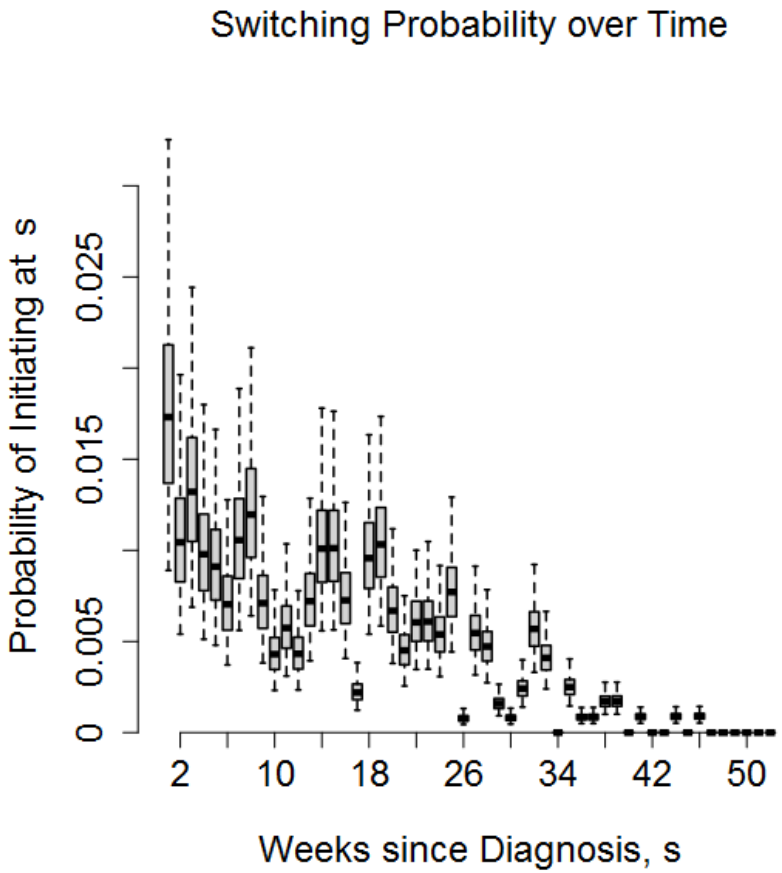
$$\beta_{0k} = \eta(1-\rho) + \rho\beta_{0k-1} + \sigma\epsilon_k$$

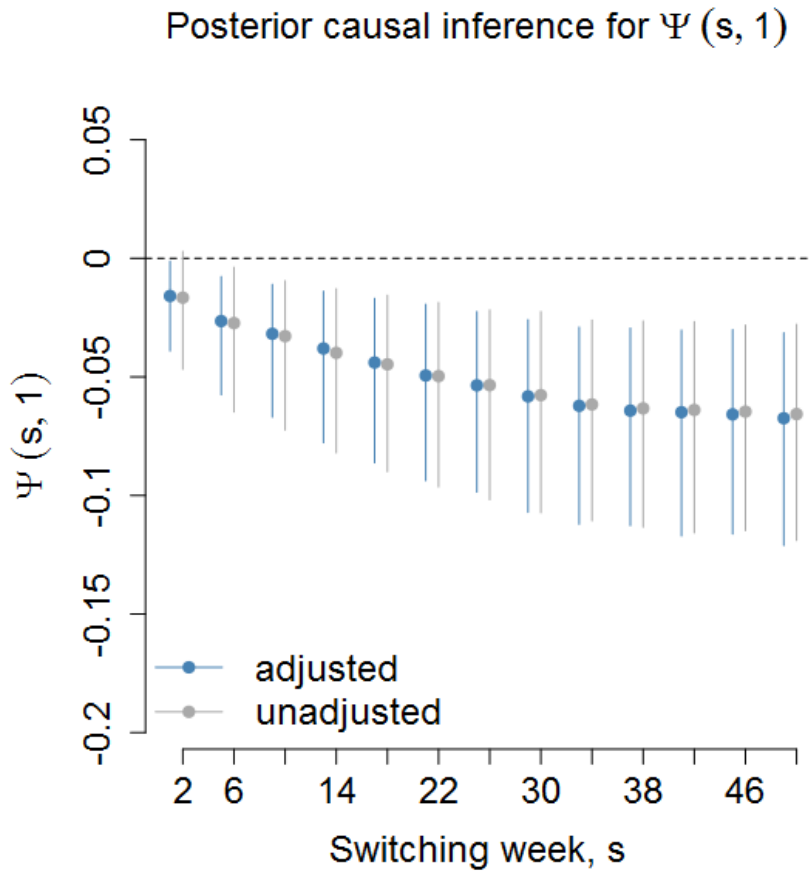
where $\epsilon_k \overset{iid}{\sim} N(0,1)$.

- Induces temporal smoothing.
- In simulations, exhibits improved MSE and credible intervals with close to nominal frequentist coverage.



Effect of Delayed Opioid Initiation on Hospitalization Rate





Takeaways and Related Work

Key takeaways:

- Event occurance and survival must be considered jointly.
- Bayesian methods can be used for principled smoothing.

Related Work:

- Oganisian A, Girard A, Steingrimsson JA, Moyo P (2024), A Bayesian framework for causal analysis of recurrent events with timing misalignment. *Biometrics*. `https://doi.org/10.1093/biomtc/ujae145`
- Oganisian A, Roy JA. (2021) A practical introduction to Bayesian estimation of causal effects: Parametric and nonparametric approaches. *Statistics in Medicine*. `https://doi.org/10.1002/sim.8761`

Web site: stablemarkets.netlify.app
X: @StableMarkets
Acknowledgements: PCORI ME-2023C1-31348.

Paper Link