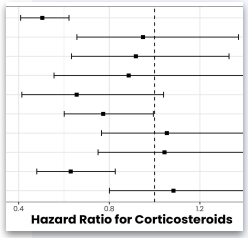
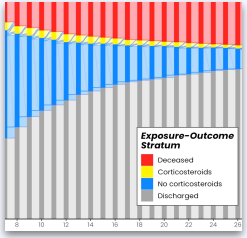
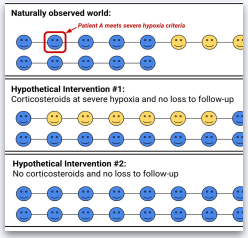
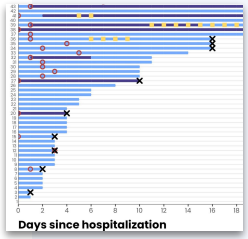
Corticosteroids in COVID-19:

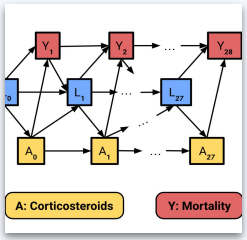
Optimizing Observational Research through Target Trial Emulations

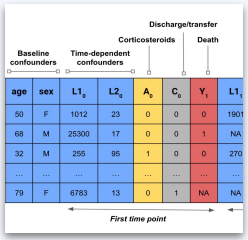












Katherine Hoffman, M.S.
Research Biostatistician
Weill Cornell Medicine

May 24, 2022

Background

- Pandemic required use of observational data for rapid decision making
- ***Corticosteroids:*** widely available drugs to combat hyperinflammation
- Dozens of observational studies are inconclusive on whether corticosteroids decrease mortality for COVID-19 patients
  - Charahom et. al meta-analysis: Odds Ratio (OR) 1.12 [95% CI 0.83-1.50]
- Multiple randomized trials (RCTs) conclude a clear benefit
  - *WHO* meta-analysis: OR 0.66 [95% CI 0.53-0.82]

A unique opportunity to compare methods for observational research against a gold-standard benchmark.

Part 1: Target Trial Emulation

- Target Trial Framework
- Non-parametric method allowing corticosteroids to be a time-varying exposure while flexibly incorporating dozens of time-dependent confounders

Part 2. Comparison to Model-first Approaches

- *Model first*: allowing the outcome type (e.g. time-to-event) to determine model type (e.g. Cox), and discussing the coefficient for treatment as a cause-and-effect relationship with the outcome

Part 1:
Target Trial Emulation

Hypothetical Target Trial

01 Eligibility

- Adults hospitalized with COVID-19 at New York Presbyterian March 1-May 15, 2020
- Not on corticosteroids for chronic use

02 Interventions

Randomized at hospitalization:

(1) If and when patient meets criteria for severe hypoxia, receive corticosteroids for 6 days

(2) No corticosteroids

03 Outcome

- 28-day mortality from time of randomization
- *Estimand*: difference in 28-day mortality rates between treatment strategies (1) and (2)

04 Data Analysis

- Hypothetical world, so no loss to follow-up
- Compute 28-day mortality rates for treatment strategies (1) and (2) and take difference

Target Trial Emulation using Observational Data

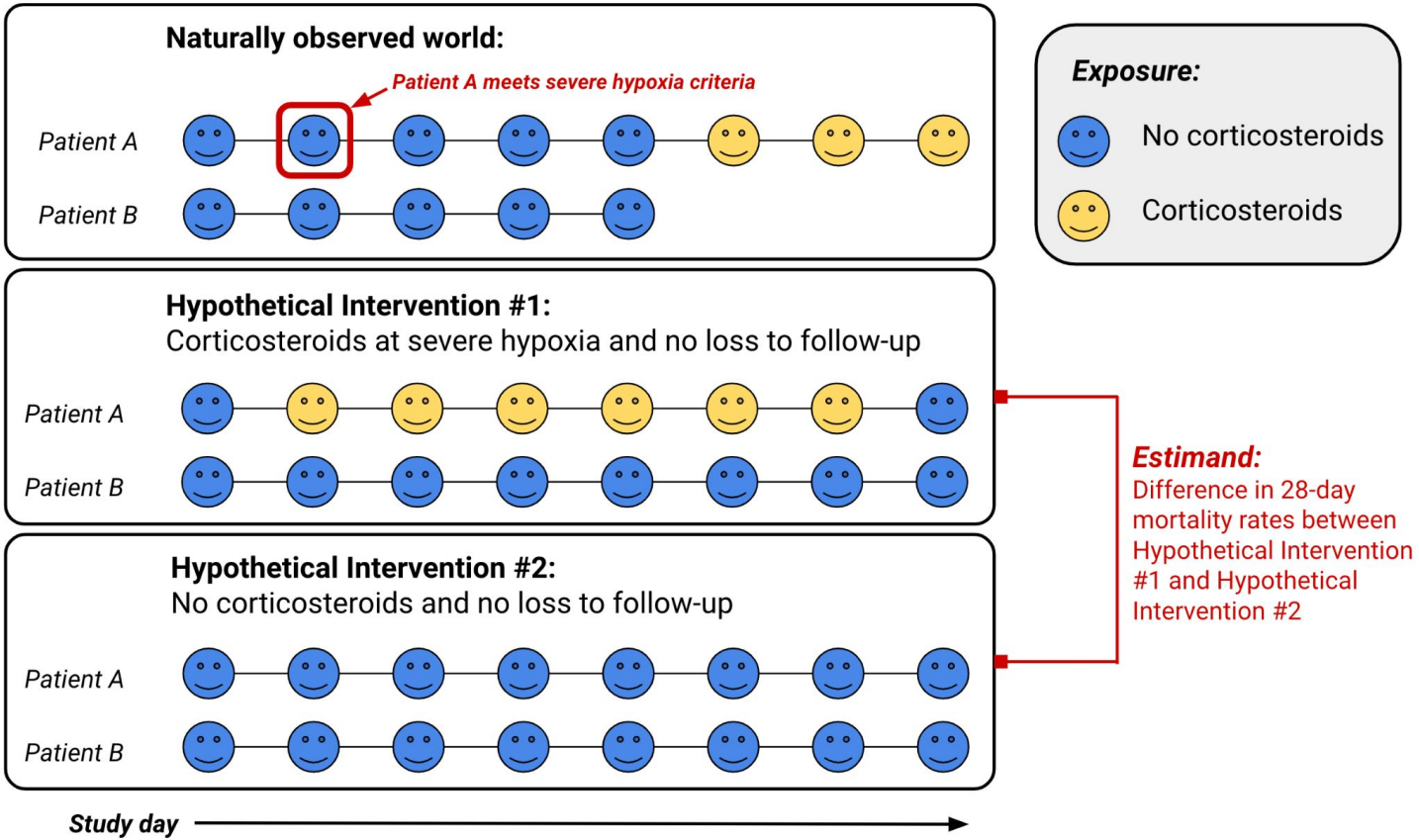
Data Source

- Retrospective, longitudinal dataset of Electronic Health Records
- All patients who met target trial's inclusion & exclusion criteria

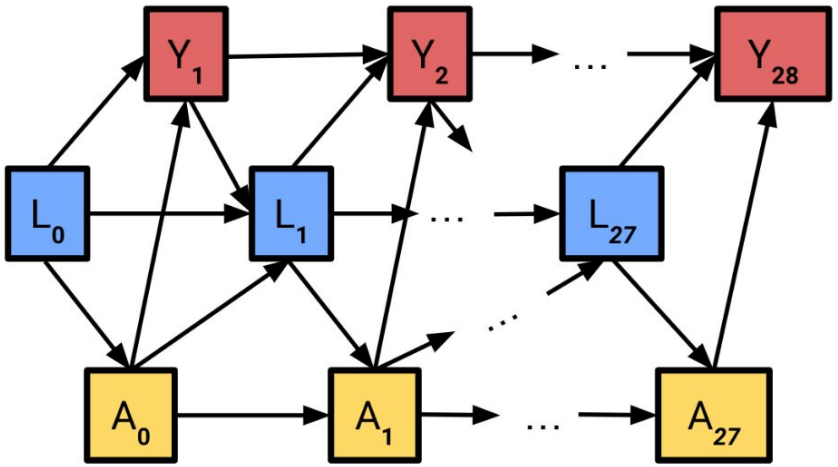
Follow up

- Patients followed for 28 days from the time of hospitalization
- Lost to follow-up by discharge or transfer to an external hospital system

Target Trial Emulation: Interventions, Outcome, and Estimand



Target Trial Emulation: Confounders



A: Corticosteroids

Y: Mortality

L: Confounders

Baseline - age, sex, race, ethnicity, Body Mass Index (BMI), comorbidities (coronary artery disease, cerebral vascular event, hypertension, Diabetes Mellitus, cirrhosis, Chronic Obstructive Pulmonary Disease, active cancer, asthma, interstitial lung disease, chronic kidney disease, immunosuppression, Human Immunodeficiency Virus, home oxygen use), mode of respiratory support upon arrival to the ED, and hospital admission location

Time-dependent - heart rate, pulse oximetry percentage, respiratory rate, temperature, blood pressure (systolic and diastolic), BUN-creatinine ratio, creatinine, neutrophils, lymphocytes, platelets, bilirubin, blood glucose, D-dimers, C-reactive protein, Activated Partial Thromboplastin time, prothrombin time, arterial partial pressure of oxygen, and arterial partial pressure of carbon dioxide, supplemental oxygen status

Target Trial Emulation: Identification Assumptions

Exchangeability

We assume a patient's covariate history contains all information about treatment, outcome, and censoring at the next time point.

Positivity

These data are from a period with high variability in provider practice, so positivity assumption is likely to be upheld.

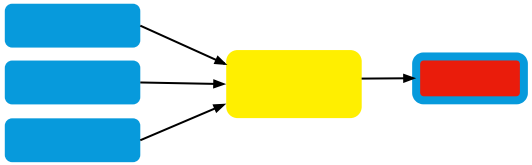
Our estimand is identified under Robin's **longitudinal g-formula**.

Target Trial Emulation: Estimation

Estimate recursive g-formula using a **sequentially doubly robust estimator**

- Can get either mechanism (intervention, outcome) correctly specified at either time point
- Allows us to use flexible machine learning regressions to estimate the mechanisms
- To learn more: "Sequential Double Robustness in Right-Censored Longitudinal Models" by Luedtke et. al (2017) and "On the multiply robust estimation of the mean of the g-functional" by Rotnitzky et. al (2017)

Superlearning/stacked regressions

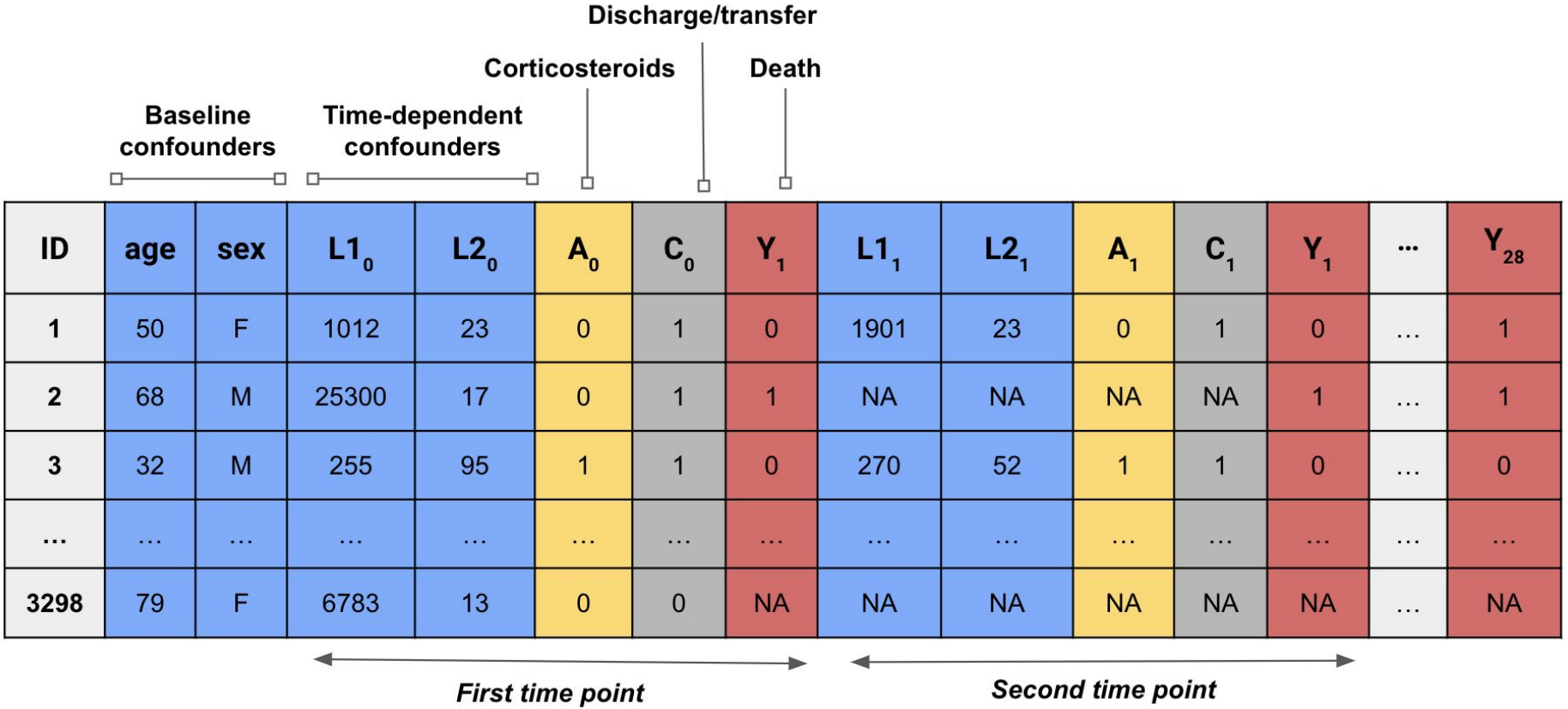


Candidate learners:

- Multivariate Adaptive Regression Splines (MARS)
- Bayesian Adaptive Regression Trees (BART)
- LASSO
- Elastic Net
- Ridge
- Mean

Cross-validation: 10 folds
Cross-fitting: 10 folds

Target Trial Emulation: Analytical File



ID	age	sex	$L1_0$	$L2_0$	A_0	C_0	Y_1	$L1_1$	$L2_1$	A_1	C_1	Y_1	...	Y_{28}
1	50	F	1012	23	0	1	0	1901	23	0	1	0	…	1
2	68	M	25300	17	0	1	1	NA	NA	NA	NA	1	…	1
3	32	M	255	95	1	1	0	270	52	1	1	0	…	0
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
3298	79	F	6783	13	0	0	NA	NA	NA	NA	NA	NA	…	NA

Target Trial Emulation: Software

- Open source R package: **lmtp**
- Define the intervention
  - Write a function that takes in current treatment and hypoxia indication and outputs counterfactual treatment
- Run function **lmtp_sdr** for both interventions
  - Specify baseline and time-varying confounders, censoring indicators, outcome, and treatment variables
  - Use **lmtp_contrast** for two intervention results to obtain final estimate
- Necessary materials to run a similar analysis publicly available on Github: **www.github.com/kathoffman/steroids-trial-emulation**

Target Trial Emulation: Results

Estimated 28-day mortality for standard of care plus corticosteroids for 6 days if and when a patient meets severe hypoxia criteria:

23% [21-24]

Estimated 28-day mortality for standard of care without corticosteroids:

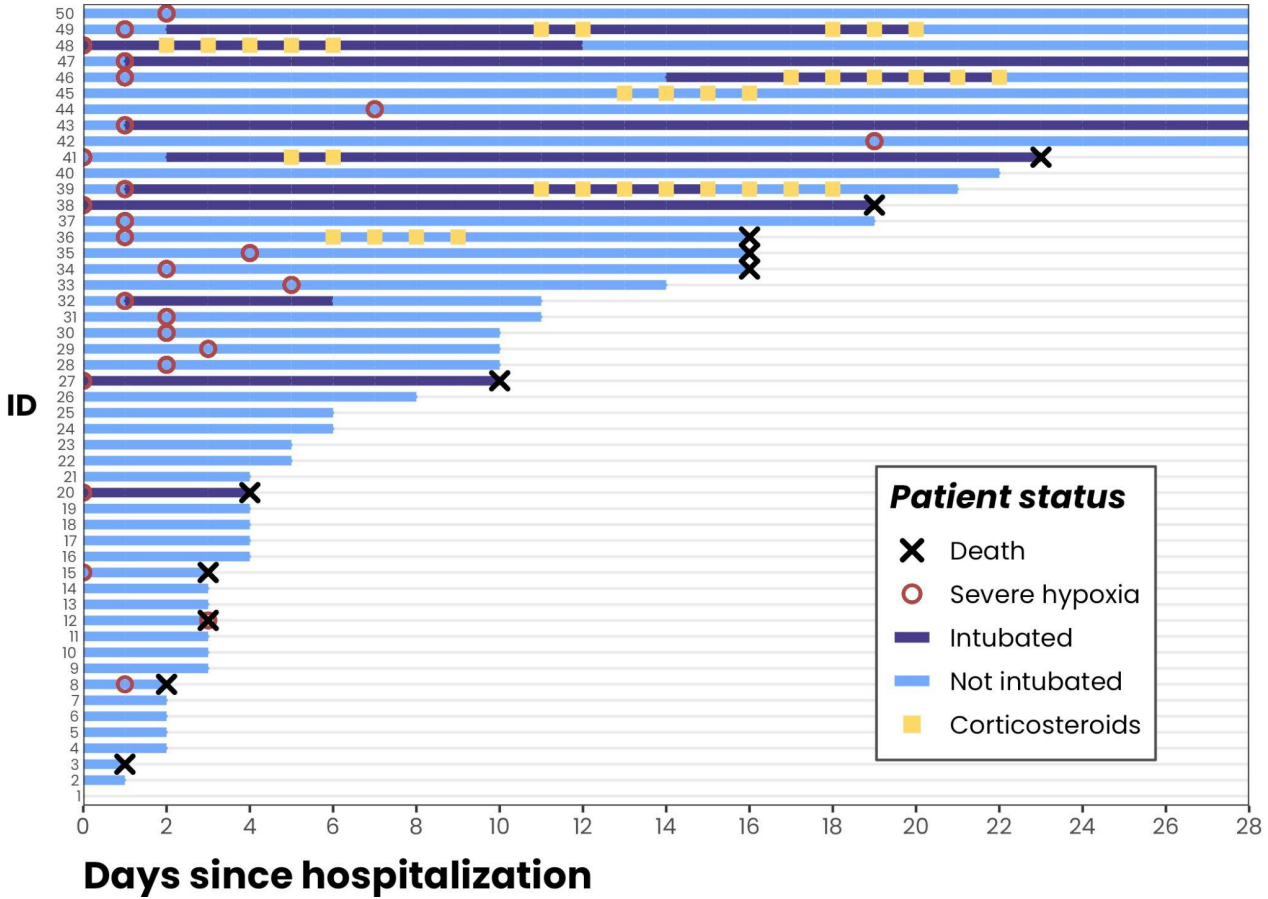
32% [31-34]

Our 6-day dynamic corticosteroids regimen would decrease 28-day mortality by **9.6% [8.8-10.4]**.

This is equivalent to an OR of 0.62, which is qualitatively identical to the WHO RCT meta-analysis of 0.66.

Part 2:
Model-first Approaches

Model-first Approach: Longitudinal Data Complexity



Model-first Approach: Methods

Regression for 28-day mortality with a point treatment entails a number of study design choices:

1. Defining a range of time relative to inclusion criteria for a patient to be considered "treated"

2. Whether to exclude patients *treated before* the inclusion criteria time

3. How to handle patients who *died during* the treatment time window

4. How to handle patients *treated after* the treatment time window

We fit Cox regressions for a point treatment using various time windows and subsequent decisions. We also fit a time-varying (TV) Cox regression.

Model-first Approach: Biases

Point-treatment Cox

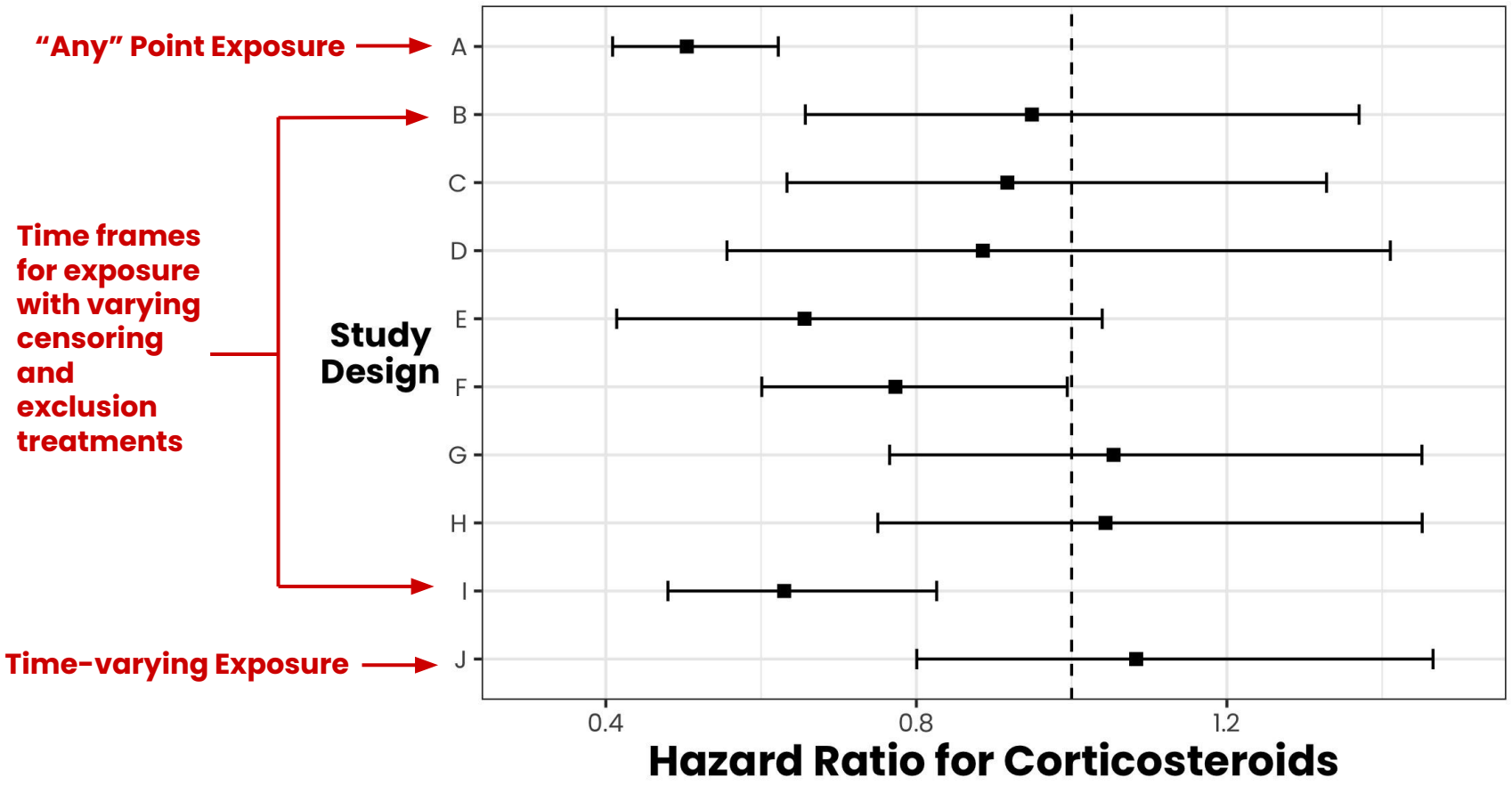
- Immortal time bias
- Cannot accommodate time-dependent confounding
- Collider bias

Time-varying Cox

- Cannot accomodate time-dependent confounding

Some of the observational research additionally include variations of propensity score matching, but this does not solve study design issues.

Model-first Approach: Results



Key Takeaways

Hypothetical Target Trial

We used a question-first approach through the Target Trial Framework and specified every component of a randomized trial we could have run.

Emulation using Observational Data

Our target trial emulation using doubly robust estimation allowed us flexibly specify an intervention, adjust for dozens of time-dependent confounders, and recovered the RCT benchmark.

Model-first Approaches

Current corticosteroids observational research suffers from various forms of bias and generally cannot recover the RCT estimates even with access to the same cohort.

Thank you

Collaborators:

Methods
Iván Díaz

Software – R pkg lmtp
Nick Williams

Clinical Expertise
Edward Schenck
Will Whalen
Di Pan
Michael Satlin

Pre-print on MedRxiv by the end of the week!

Twitter: @kat_hoffman_

Blog: www.khstats.com

Email: kathoffman.stats@gmail.com