Corticosteroids in COVID-19: Optimizing Observational Research through Target Trial Emulations

# **Supplemental Materials**

## **Methods**

### **Confounders**

In cases of multiple laboratory results or vital signs in a 24-hour period, the clinically worst value was used. Missing categorical baseline variables were handled using an indicator for missingness. Missing continuous baseline variables were handled using mean imputation with an indicator for missingness added as a covariate. Missing laboratory results and vital signs were all continuous and handled using a combination of last observation carried forward or mean imputation when no previous measures were recorded. An indicator for missing data was used to account for potentially informative time-dependent missingness mechanisms. All potential confounders were determined through input from clinicians.

### **Estimand**

The g-computation formula can be conceptualized as a sequential regression of the outcome as follows(48). First, a model for the probability of mortality a day 28 is estimated amongst patients who have not been lost to follow-up, with predictors given by all the data prior to day 28. This model fit is then used to produce predictions of the probability of mortality for all patients had the hypothetical intervention of interest been implemented at day 28. This probability is then used as a pseudo-outcome in a regression model where the predictors are all variables measured up to day 27. This model is then used to obtain new predictions for all patients had the hypothetical intervention of interest been implemented at day 27. This process is iterated until day 1, at which point the predictions are averaged. This average is the estimated mortality rate under the hypothetical treatment rule.

### **Data analysis plan**

The SDR estimator uses the probability of receiving the intervention, the probability of experiencing the outcome, and the probability of censoring at each time point to estimate the final survival probability under a given intervention. For longitudinal studies, the SDR estimator estimates these probabilities iteratively at discrete time points. At each time point, only the intervention or outcome model need be correctly specified for the final estimate to be consistent(21). We fit each model using an ensemble machine learning via the super learner algorithm(25, 26). The super learner algorithm creates a weighted combination of algorithms in a user-supplied library to obtain a final prediction that is guaranteed to outperform each algorithm in the library as sample size grows(26).

We use a super learner library containing candidate learners of multivariate adaptive regression splines (MARS) with varying parameters, Bayesian Additive Regression Trees (BART), penalized linear regression with penalties of 0, 0.5, and 1 (Ridge, Elastic Net, and LASSO), and an interceptonly (mean) model. To fit these regressions, we assume that outcomes at day *t* only depend on baseline confounders and time-varying covariates measured in the previous two days.

### **Analytical file**

Supplemental Figure 2 shows an illustrated example of the analytical file. Time-dependent confounders, censoring indicators, and outcome indicators are binned in 24-hour windows from time of randomization. The final analytical file contains one row per patient with one column for each baseline confounder and one column for each time-dependent confounder, outcome, and censoring indicator at all 28 time points.



Supplemental Figure 1: A random sample of 50 patients’ observed hospital courses. Each line represents one patient, and the color of the line represents that patient’s intubation status (a time-dependent confounder to corticosteroids administration and death). A red circle indicates the time of severe hypoxia. Yellow squares indicate steroids administration. We see that patients received corticosteroids at various times and durations throughout their hospitalization course and in relation to meeting the criteria of severe hypoxia. This is ideal for experimentation, but makes designing a point-treatment study extremely difficult.