Improving upon Intention-to-Treat Analysis when Clinical Trials become Open-Label

Samantha R. Cook Donald B. Rubin Harvard University
cook@stat.harvard.edu

In a recent FDA trial, the drug under investigation became commercially available before the end of the trial. Patients in the trial therefore had the option of going off trial protocol and obtaining the commercially available active therapy. When patients randomized to placebo switch to commercially available therapy, they cease to be controls in the usual sense: All measurements after a placebo control switches to active therapy are missing. We propose a method to impute placebo controls' missing outcomes, as if they had stayed on placebo. This involves first, fitting a Bayesian hierarchical regression model to data from untreated historical patients; second, incorporating information learned from the historical patients into a similar model for placebo controls; and third, using this model and observed on-protocol data to impute missing values for placebo controls who switched to active therapy. Using multiple imputation will allow for improved estimation of the treatment effect and a more powerful test of the drug's effectiveness.

Project Overview

- Treatment in Phase 4 trial became commercially available early.
- Some placebo patients began taking active treatment.
- Outcomes under placebo are missing after this point – want to impute them.
- Imputing based only on observed data would rely heavily on extrapolation.
- Develop model for disease progression using historical data.
- Incorporate information from historical patients into model for randomized patients.
- Use this model and on-protocol data to impute missing values.

Fabry disease

- Rare genetic disease caused by a defect on the X chromosome. Results in inability to produce enough of the enzyme α-GAL, which leads to creatinine buildup in the blood.
- Symptoms involve pain, impaired sweating, skin rashes, and can lead to serious problems with the heart and kidneys.
- Previous medical practice focused on treating the symptoms of the disease, ranging from rest and changes in diet to kidney dialysis or coronary bypass surgery.
- Fabrazyme: first drug approved for treating Fabry disease.
- Phase 4 trial of Fabrazyme recently concluded–waiting for unblinded data.

Phase 4 Trial

- 72 patients: 1/3 assigned placebo, 2/3 assigned Fabrazyme.
- Enrollment criteria: Mild to moderate renal disease.
- Primary outcome: Clinical event (death, dialysis, stroke).
- Monthly measurements of serum creatinine: Signifi cant increase in serum creatinine considered a clinical event.
- Primary analysis: Log rank test.
- Goal: Impute missing placebo (serum creatinine) outcomes and perform log rank test.

Historical Data Base

- Assembled from medical records and patient registries.
- 15+ years of data.
- 66 patients.
- 420 observations.
- Measured serum creatinine as well as background covariates
  Age, weight, gender, blood type.

- Estimated propensity scores: Propensity to be in current randomized study vs. historical data set.

⋆ Chose sample of historical patients who looked, at baseline, most like the set of patients in current clinical trial.

Modeling Serum Creatinine

- Serum creatinine levels remain relatively constant until an unobserved change-point, after which serum creatinine begins to increase. Increase generally continues until some external intervention, e.g., kidney dialysis.
- Rates of increase vary across people.
- Patients post-baseline are presumably in phase of increasing serum creatinine. Inclusion criteria for Phase 4 study required that patients have elevated serum creatinine or signs of renal disease.
- Inverse serum creatinine more linear, standard outcome considered in Fabry literature.
- Model: Within each patient, inverse serum creatinine is non-increasing, rates of change vary across people.

Data Model

Sampling Distribution

$$y_{ij} \sim N(\mu_i - \exp(\beta_i)t_{ij} - \exp(\gamma_i)t_{ij}^2,\ \sigma^2)$$

$$i = 1, \ldots, 66 \quad j = 1, \ldots, n_i$$

Prior Distribution

$$(\mu_i, \beta_i, \gamma_i) \sim MVN(\boldsymbol{Z}_i\boldsymbol{\xi},\ \boldsymbol{\Psi})$$

$$\boldsymbol{Z}_i = (1, PS_i, PS_i^2)$$

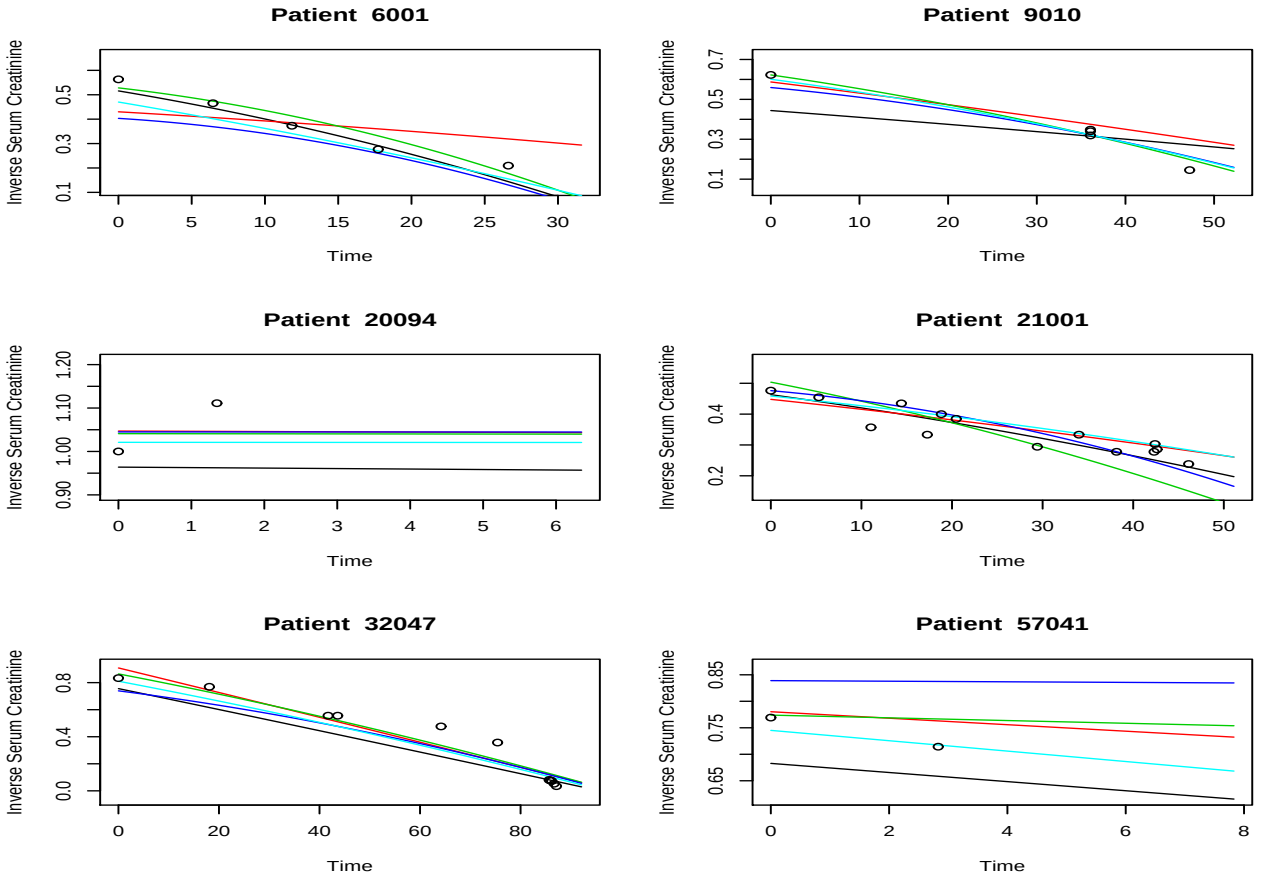
$$i = 1, \ldots, 66$$

Vague prior distribution on σ^2, $\boldsymbol{\xi}$, and $\boldsymbol{\Psi}$.

$\boldsymbol{\eta} \equiv$ parameters in $(\boldsymbol{\xi}, \boldsymbol{\Psi})$ governing marginal prior distribution of γ
(nonlinear long-term trends)

Predicted Serum Creatinine Progression

Mean inverse serum creatinine based on posterior draws of model parameters



Incorporating Information from Historical Database into Model for Randomized Controls

- Fit model to historical patients' data.
- Use posterior distribution of η from historical patients as prior distribution in similar model for placebo controls.
- Long-term trends will be well estimated for historical patients but less well estimated for randomized patients, due to shorter follow-up.
- Other parameters should be well estimated from randomized controls' data, due to frequent follow-up.
- Incorporate prior information only where it's really necessary.

Imputing Missing Data

- Fit model to randomized controls' observed data.
- Obtain draws of model parameters for each subject.
- Draw missing values given these parameters and covariates, i.e., given the ℓth parameter draw,

$$y_{ij^*}^{(\ell)} \sim N(\mu_i^{(\ell)} - \exp(\beta_i^{(\ell)})t_{ij^*} - \exp(\gamma_i^{(\ell)})t_{ij^*}^2,\ \sigma_i^{2^{(\ell)}})$$

  j^* = time points with missing observations.

- Event Time: First time point where imputed serum creatinine increased by 33% or more.
- Log rank test statistic can be computed for each imputed data set and the results combined using standard rules.
- Method should be conservative: The only event that can be imputed for placebo controls is 33% increased serum creatinine.

Model Validation

- Simulation studies address primary analysis
- First examine validity of log rank test with complete data
  - Generate data under null hypothesis that there is no difference in event rate between Treatment and Placebo groups
  - Proportion of data sets where null hypothesis is rejected is the true signifi cance level of the test
- Test procedure with missing data
  - Impose the pattern of missing data seen in randomized study
  - Impute missing data using model developed here
  - Level of test should be close to the level with complete data

Validation Simulation

- Generate data under null hypothesis
- "Shifted" random walk:
  $SC_{ij} \sim N(k * SC_{i,j-1},\ \sigma^2)$
  $y_{ij} = SC_{ij}^{-1}$
- SC_{i0} = baseline observation
- $k = 1.01$, $\sigma = .05$
- 8 patients dropped out

Validation Results

Test Level:

- Complete data level = .062
- MI level = .065
- p = .686

Power:

Generate Fabrazyme patients' data under shifted random walk with $k^* < 1.01$

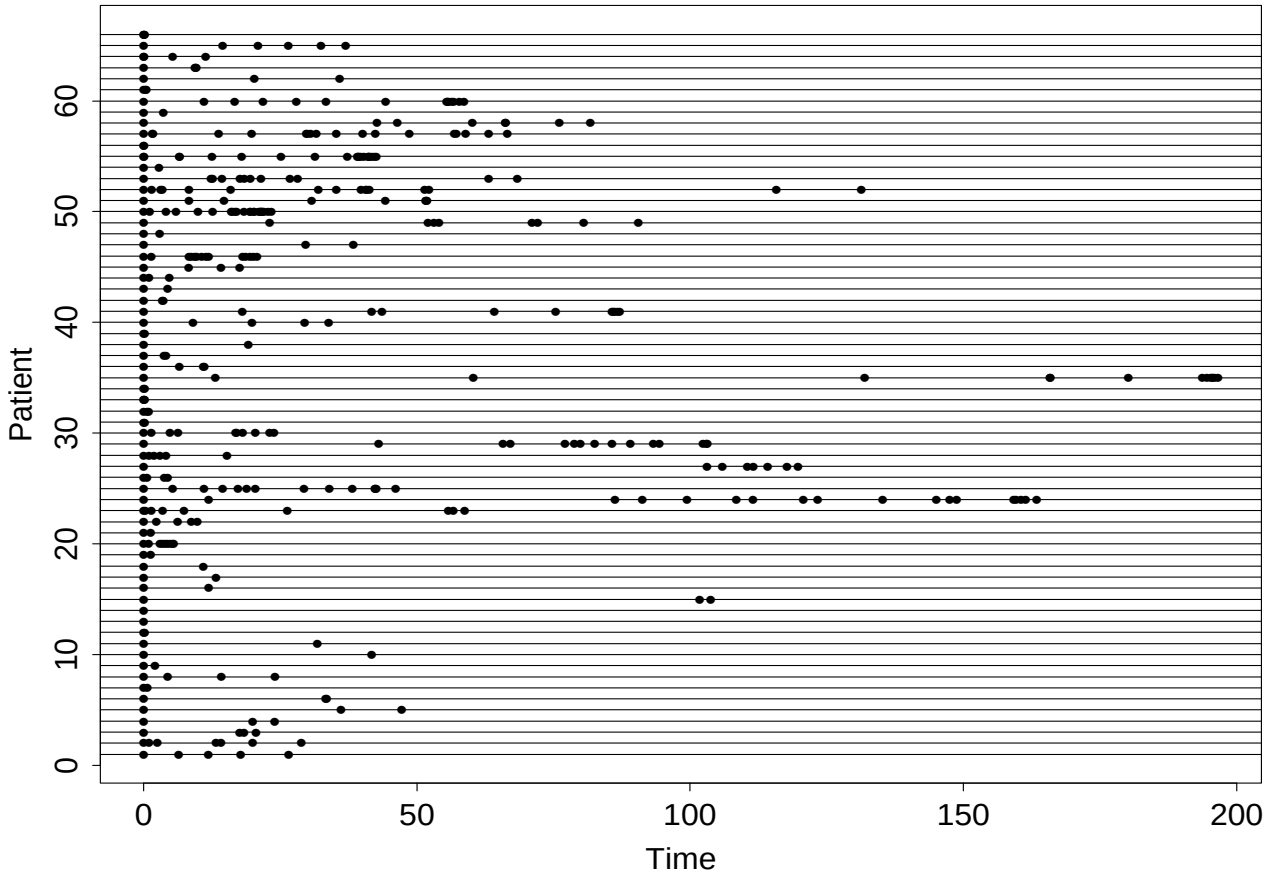
- $k^* = 1.008$ Censored data power = .19 MI power = .30
- $k^* = 1.004$ Censored data power = .77 MI power = .88
- $k^* = 1$ Observed data power = .90 MI power = .93

Conclusions and Future Work

- Flexible class of models for disease progression:
  - Straightforward to incorporate constraints on progression
  - Accommodates both time-varying and person-level covariates
  - Accommodates irregular timing of measurements
- By using historical data only as part of the prior distribution, we use outside information only where it is most necessary
- By imputing missing data we increase power while maintaining test validity
- Next step: Analyze trial data
- Sensitivity analyses

Historical Patients' Data

Timing of Historical Patients' Observations



Multiple Imputation

- Repeatedly fi ll in missing values with plausible values.
- Complete data: $Y = \{Y_{obs}, Y_{mis}\}$.
- Imputation Model: $p(Y|\theta)p(\theta)$.
- $Y_{mis} \sim p(Y_{mis}|Y_{obs}) = \int p(Y_{mis}, \theta|Y_{obs})d\theta =$
  $\int p(Y_{mis}|Y_{obs}, \theta)p(\theta|Y_{obs})d\theta$.
- Repeatedly generate imputations $Y_{mis,1}, Y_{mis,2}, ..., Y_{mis,m}$.
- Analyze each imputed data set as if all imputed values were observed values.
- Combine complete-data inferences into single inference using simple rules.
- Rubin (1976), Schafer (1997), Little and Rubin (2002)