τ -Inflated Beta Regression Model for Estimating τ -Restricted Means and Event-Free Probabilities for Censored Time-to-Event Data.

Joint work with Yizhuo Wang

Source Material:

https://pmc.ncbi.nlm.nih.gov/articles/PMC11604032

https://github.com/yezow/--Inflated-Beta-Regression-Model-for-Estimating--RMST (Manuscript Github Site)

Ann Arbor ASA Meeting Github Site: https://github.com/yezow/tau-IBR_ASA

Goals for this presentation

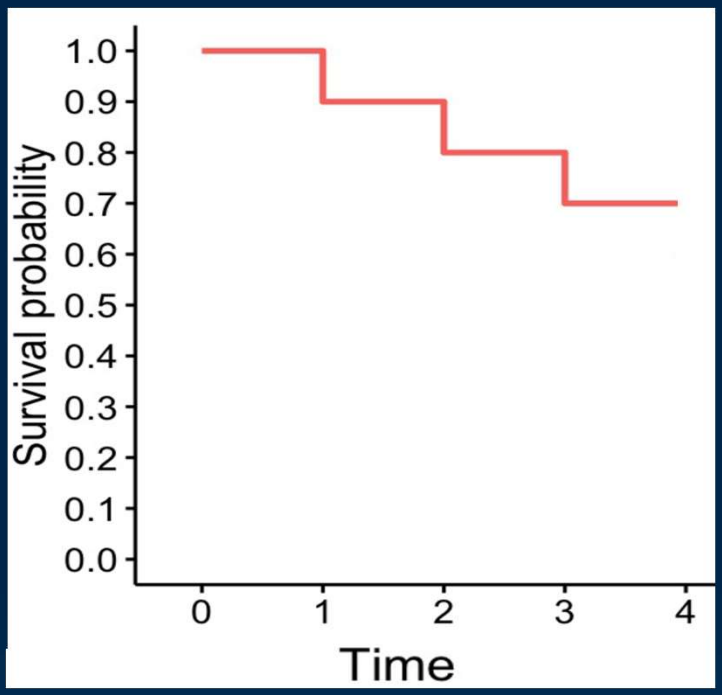
- Review existing τ- restricted mean survival time (τ-RMST) regression model for τ-restricted event times, $\min(\tau, T_i)$ (including R code for estimating these models)
- Motivate τ- Inflated Regression (τ-IBR) Model
  - Fundamental identity:

$$\min(\tau, T_i) = \tau\, I(T_i \geq \tau) + T_i\, \mathrm{I}(T_i < \tau)$$

- Describe the τ-IBR Model
  - Model specification
  - τ-RMST estimation
- Intuition gained through simulation study
- Example analyses with code

τ-Restricted Event Times, $\min(\tau, T_i)$

- We work with these event times nearly always
  - Finite follow-up periods common
  - Suppose $\tau =$ largest observed follow-up time



- For $t < \tau$, the Kaplan-Meier (KM) estimate for $P(T_i > t)$ is equivalent to the KM estimate for $P(\min(\tau, T_i) > t)$
- Area under KM curve up to τ estimates $E[\min(\tau, T_i)]$, or the τ-RMST

τ-RMST Regression Model

- Most well known regression model for analysis of τ-restricted event times, $\min(\tau, T_i)$, introduced by Andersen, Hansen and Klein (2004)

$$E[\min(\tau, T_i)] = \beta_0 + \beta_1 Z_{1i} + \cdots + \beta_p Z_{pi}$$

  - Estimates the expected event-free years during τ years of follow-up from covariates
  - Paper introduced pseudo-observations, $PO_i, i = 1, \ldots, n$, as an approach for estimating model parameters with censored outcomes
  - Pseudo-observations, $PO_i, i = 1, \ldots, n$, stand in for $\min(\tau, T_i), \mathrm{i} = 1, \ldots, n$, in the model fitting process
  - Standard generalized linear model (GLM) software obtains parameter estimates
  - Identity and log links both in common use

Pseudo-Observations for $\min(\tau, T_i)$

- For the i^{th} individual in the dataset, the pseudo-observation, PO_i, corresponding to their $\min(\tau, T_i)$ outcome is:

$$PO_i = n * RMST - (n-1) * RMST^{(-i)},$$

where $RMST$ is the τ-RMST estimate using the whole cohort and $RMST^{(-i)}$ is the τ-RMST estimate using everyone in the cohort except patient i.

- Andersen, Hansen and Klein showed that PO_i has the same conditional expectation as the (potentially censored) $\min(\tau, T_i)$ outcome

- That is, we can estimate parameters from the desired model,

$$\mathrm{E}\{\min(\tau, T_i)\} = \beta_0 + \beta_1 Z_{1i} + \cdots + \beta_p Z_{pi},$$

by fitting the model

$$\mathrm{E}\{PO_i\} = \beta_0 + \beta_1 Z_{1i} + \cdots + \beta_p Z_{pi}$$

- A 'jackknife methodology' approach

Maximum Possible τ When Using Pseudo-Observations

$$PO_i = n * RMST - (n-1) * RMST^{(-i)},$$

where $RMST$ is the τ-RMST based on whole cohort and $RMST^{(-i)}$ is the τ-RMST based on whole cohort without patient i.

- τ must be a valid choice for estimating $RMST^{(-i)}$ for $i = 1, \ldots, n$.
  - At least one of the two largest event times is included in $RMST^{(-i)}$. So, when figuring out the maximum possible τ, you only need to think about these two values.
  - No restriction on τ if largest two event times are both deaths.
  - If both of the largest two event times are censored, then the maximum possible τ is the smaller of these two values.
  - If only one of the largest two event times is censored, then the maximum possible τ is this value.

Example: Azithromycin for Prevention of Chronic Obstructive Pulmonary Disease (COPD) Exacerbations Trial

- Patients were randomized to azithromycin or placebo.
- Followed for the primary endpoint of time-to-first acute exacerbation during the subsequent year.
- A 1-year-RMST model estimates acute exacerbation-free days over the follow-up year by treatment group and other patient characteristics
  - For $\tau = 1$, equivalent to interpret parameters as (absolute) increase in the percent exacerbation-free time during the 1-unit follow-up period corresponding to a 1-unit increase in the predictor
  - Advantage to rescaling outcomes from $(0, \tau]$ scale to (0,1] scale when interpreting results

Example: Azithromycin for Prevention of Chronic Obstructive Pulmonary Disease (COPD) Exacerbations Trial

- Variables in ex_data_example.csv:
  - X= observed follow-up days on study
  - Delta=1 if follow-up time, X, corresponds to an exacerbation, 0 otherwise
  - trt = 1 if azithromycin, 0 otherwise
  - age_10_new = age in decades, centered around 65 years
  - gender=1 if male, 0 otherwise
  - fev1_10_new = FEV1 per 10% predicted, centered around 40% of predicted
  - nowsmk=1 if current smoker, 0 otherwise
  - Site indicator variables: latitude_34, latitude_38, latitude_39, latitude_40, latitude_42, latitude_44

R Code for τ-RMST Model (Read in Data)

Data_cens<- read.csv("ex_data_example.csv")

#Key data shown below (except for site indicator variables)

X	Delta	trt	age_10_new	gender	fev1_10_new	nowsmk
84	1	0	0.08145815	1	-0.34759831	0
360	0	0	-1.31854185	1	1.23404754	1
140	1	1	1.18145815	1	0.92965445	0
351	1	0	-0.71854185	1	-2.13845969	0
66	0	0	-1.01854185	1	-1.40125483	0
364	0	0	-0.21854185	1	1.56789132	0
26	1	0	-0.91854185	0	2.07459421	0
361	0	1	0.28145815	1	-1.24743238	0
129	1	1	-0.71854185	1	-1.45947188	1

R Code for τ-RMST Model (Create Pseudo Observations)

```
library(pseudo)
set.seed(1234)
tau= 365

#Create pseudo-observations
pseudo <- pseudomean(Data_cens$X,
          Data_cens$Delta,
          tmax = tau)

#Showing outcomes on year scale:
head(cbind(X=Data_cens$X/tau,
          Delta=Data_cens$Delta,
          Pseudo=pseudo/tau), 10)
```

X	Delta	Pseudo
0.230136986	1	0.222603283
0.986301370	0	1.010197920
0.383561644	1	0.371845599
0.961643836	1	0.965209584
0.180821918	0	0.753752513
0.997260274	0	1.010393102
0.071232877	1	0.070089112
0.989041096	0	1.010393102
0.353424658	1	0.341515767

⋮

Pseudo-observations have the same conditional expectation as uncensored data values would have, but don't necessarily look like the uncensored outcomes of interest.

R Code for τ-RMST Model (Fit Model)

```
library(stats)
pseudo_fit <- lm(pseudo/tau~
          trt+
          age_10_new+
          gender+
          fev1_10_new+
          nowsmk+
          latitude_38+latitude_39+latitude_40+
          latitude_42+latitude_44,data=Data_cens)
summary(pseudo_fit)
```

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.597749   0.032569  18.354  < 2e-16 ***
trt          0.091471   0.022792   4.013 6.39e-05 ***
age_10_new   0.035193   0.014373   2.449 0.014499 *
gender       0.079823   0.023578   3.385 0.000736 ***
fev1_10_new  0.023138   0.007607   3.042 0.002410 **
nowsmk       0.013907   0.029807   0.467 0.640900
latitude_38 -0.096305   0.048329  -1.993 0.046541 *
latitude_39 -0.170618   0.049043  -3.479 0.000523 ***
latitude_40 -0.102577   0.032800  -3.127 0.001810 **
latitude_42 -0.176630   0.035893  -4.921 9.92e-07 ***
latitude_44 -0.205978   0.041692  -4.940 9.00e-07 ***
```

Manuscript Friendly Results With Coefficients, 95% CIs, P-values

```
coefs<-pseudo_fit$coefficients
se_coefs<-coef(summary(pseudo_fit))[, 2]
pvals<-as.numeric(
          summary(pseudo_fit)$coefficients[,4])

CI_RMST_lower<- coefs -1.96*se_coefs
CI_RMST_upper<- coefs +1.96*se_coefs
RMST_model_results<-rbind((coefs)[2:6],
                    (CI_RMST_lower)[2:6],
                    (CI_RMST_upper)[2:6],
                    (pvals)[2:6])
round(RMST_model_results,digits=3)
```

```
           trt age_10_new gender fev1_10_new nowsmk
coef_tau 0.091      0.035  0.080       0.023  0.014
lower_CI 0.047      0.007  0.034       0.008 -0.045
upper_CI 0.136      0.063  0.126       0.038  0.072
p_value  0.000      0.014  0.001       0.002  0.641
```

Example: Azithromycin for Prevention of Chronic Obstructive Pulmonary Disease (COPD) Exacerbations Trial

	Azithromycin (vs. Placebo)	Age (per 10 Years)	Male (vs. Female)	FEV_1 (per 10% Predicted)	Current Smoker (vs. Ex)
τ-RMST Model					
Coef/τ	0.091	0.035	0.080	0.023	0.014
95% CI	(0.047, 0.136)	(0.007, 0.063)	(0.034, 0.126)	(0.008, 0.038)	(-0.045, 0.072)
P-value	<0.001	0.014	<0.001	0.002	0.641

Results are additionally adjusted for the study site (data not shown)

- During the year on study, the percentage of time spent exacerbation-free was estimated to be 9.1% longer when taking azithromycin versus placebo (95% CI 4.7-13.6% longer; p<0.001), when adjusted for age, sex, FEV1 % predicted, current smoking status and site.

Motivation for τ- Inflated Beta Regression (τ-IBR) Model

- Fundamental identity:

$$\min(\tau, T_i) = \tau\, I(T_i \geq \tau) + T_i\, \mathrm{I}(T_i < \tau)$$

is a mixture random variable that can be rewritten in terms of

$$B_i = I(T_i \geq \tau) \sim \text{Bernoulli}(\pi_i) \text{ with mean, } \pi_i = \mathrm{P}(T_i \geq \tau)$$

and

$$Y_i = \frac{T_i}{\tau} \text{ given } T_i < \tau,$$

which we will assume follows a beta$[\mu_i \nu, (1-\mu_i)\nu]$ distribution with mean, μ_i.

- In terms of these random variables,

$$\min(\tau, T_i) = \tau[B_i + Y_i(1 - B_i)]$$

- Suggests modeling of B_i and Y_i as alternative to τ-RMST model

Components of τ-IBR Model

- Logistic regression model for $B_i = I(T_i \geq \tau)$:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 Z_{\pi 1 i} + \cdots + \beta_{p_1} Z_{\pi p_1 i}$$

What individuals tend to be event-free during follow-up?

- Beta regression model for $Y_i = \frac{T_i}{\tau}$ given $T_i < \tau$:

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \alpha_0 + \alpha_1 Z_{\mu 1 i} + \cdots + \alpha_{p_1} Z_{\mu p_1 i}$$

Amongst those who have the outcome during follow-up, what individuals have a higher percent of event-free time?

- Framework achieves a couple of important goals:
  - Directly addresses the point mass at τ if goal is τ-RMST estimation based on covariates
    - $\tau\text{-RMST}_i = \tau[\mu_i(1-\pi_i) + \pi_i]$
  - Addresses the possibility that different covariates may play a role in the B_i and Y_i regression models

Interpreting Coefficients from τ-IBR Model

- Logistic regression model for $B_i = I(T_i \geq \tau)$:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 Z_{\pi 1 i} + \cdots + \beta_{p_1} Z_{\pi p_1 i}$$

  - For a 1-unit increase of $Z_{\pi k i}$ from z to z + 1, with all other covariates fixed, the odds ratio for $B_i = I(T_i \geq \tau) = 1$ becomes e^{β_k}

- Beta regression model for $Y_i = \frac{T_i}{\tau}$ given $T_i < \tau$:

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \alpha_0 + \alpha_1 Z_{\mu 1 i} + \cdots + \alpha_{p_1} Z_{\mu p_1 i}$$

  - For a 1-unit increase of $Z_{\mu k i}$ from z to z + 1, with all other covariates set to zero, the fold change for $\mu_i = \frac{1}{\tau}\mathrm{E}(T_i \mid T_i < \tau)$ becomes

$$\frac{e^{\alpha_k}(1 + e^{\alpha_0 + z\alpha_k})}{(1 + e^{\alpha_0 + \alpha_k + z\alpha_k})}$$

  - Helpful to set all covariate reference values to zero

Censoring Addressed by Fitting Models Using One of Two Approaches

- Approach 1: Estimation-maximization (EM)
- Approach 2: Inverse-transform multiple imputation (MI) to fill in missing information for censored outcomes
  - Defines risk sets of similar individuals to censored individual in terms of (μ_i, π_i) and nonparametrically samples missing residual lifetime from risk set's KM estimate
  - *M* complete datasets obtained via MI algorithm
  - Results combined across imputed datasets via Rubin's rules (Little and Rubin, 1987. Analysis with missing data. JohnWiley & Sons.)
- Technical details omitted, but both EM and MI approaches work well
  - MI approach allows for additional analyses, plots

Learning Through Simulation

- Setting:
  - n=1000, $\tau = 30$
  - Independent covariates, where $Z_{1i} \sim$ Unif(0,1), $Z_{2i} \sim$ Bernoulli(0.7), $Z_{3i} \sim$ Unif(0,1)
- Logistic regression model for B_i:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -1.0 + 1.0Z_{1i} + 2.0Z_{2i} - 1.5Z_{3i}$$

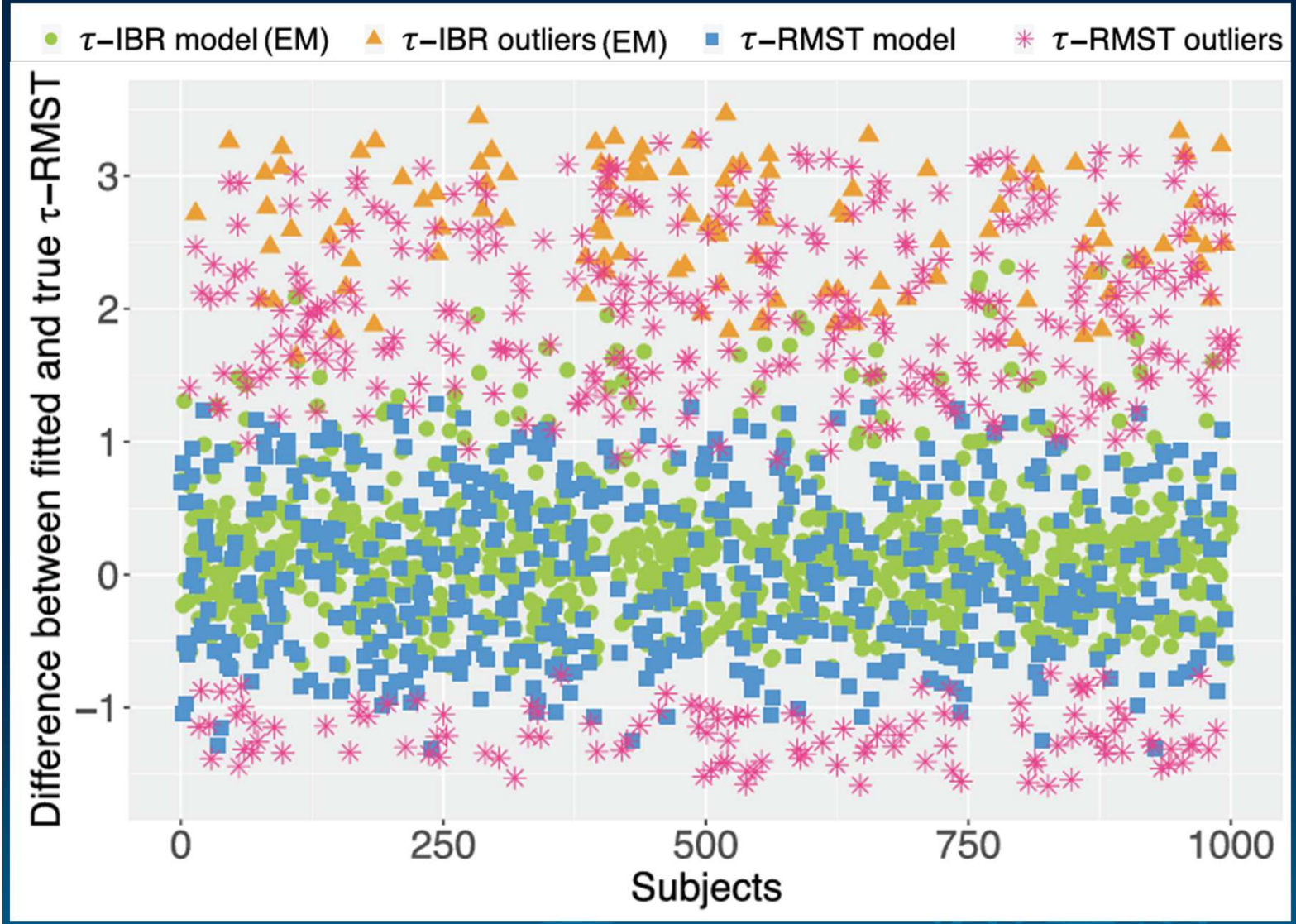
- Beta regression model for Y_i:

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = -2.0 + 1.2Z_{1i} + 2.0Z_{2i}$$

- Approximately 52% of simulated individuals would have $\min(30, T_i) = 30$, if uncensored
- Independent censoring generated so that approximately 30% censoring of events prior to 30.

Key Simulation Findings

- Both τ-IBR and standard τ-RMST models gave unbiased estimates for $\mathrm{E}\{\min(\tau, T_i)\}$
- However, τ-IBR estimates were more precise and had better coverage probabilities than τ-RMST estimates

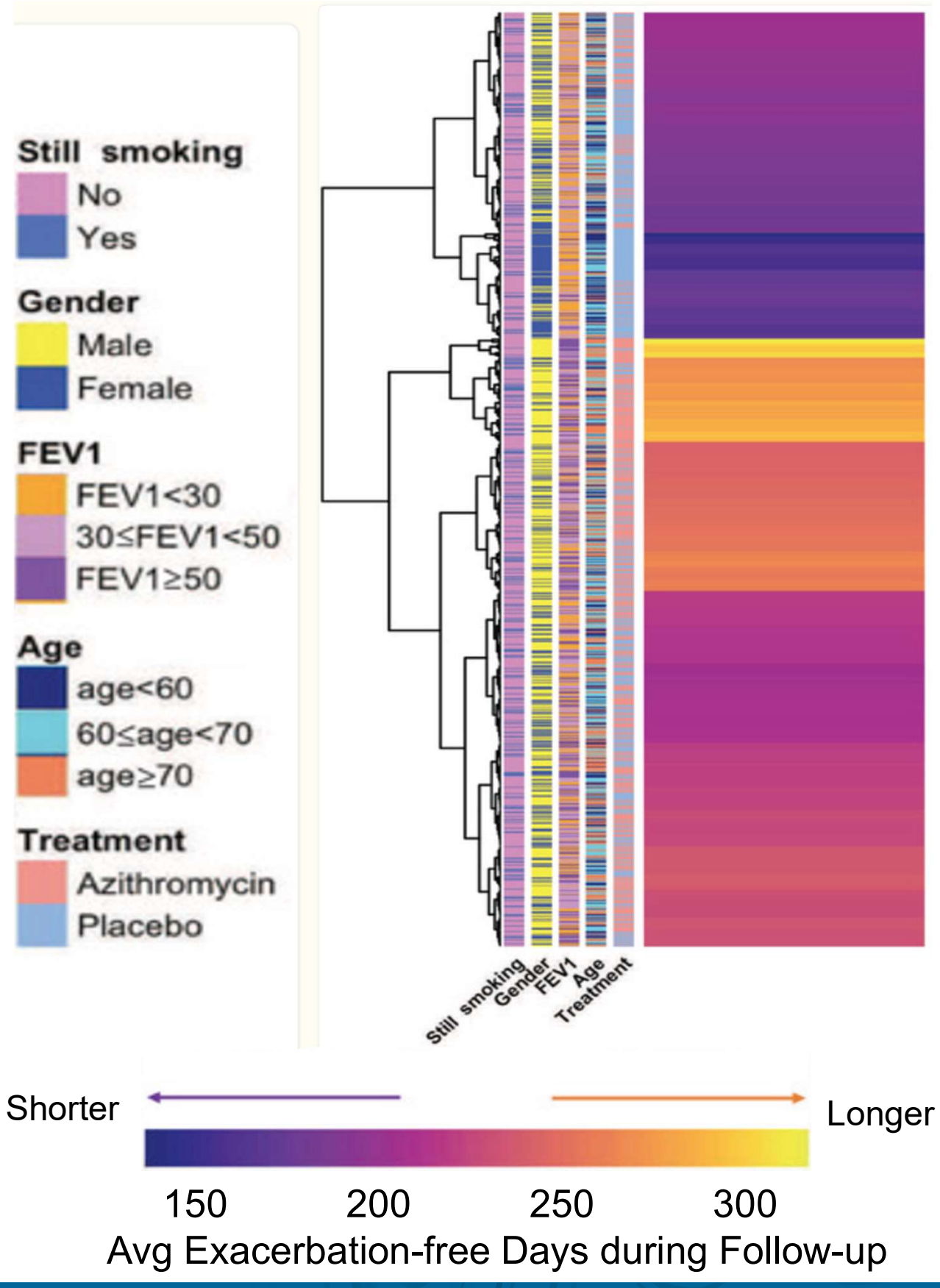


- Both models perform similarly if there are no important covariates in the logistic component of the τ-IBR model

τ-IBR Example: Azithromycin for Prevention of Chronic Obstructive Pulmonary Disease (COPD) Exacerbations Trial

- One feature of the τ-IBR MI algorithm is that we obtain multiple completed datasets of $\min(\tau, T_i)$ outcomes.
- This allows us to plot data using methods available for uncensored outcomes (see heatmap).

Min(365 day, T_i) Values Averaged Across 10 MI Datasets Using τ-IBR Method)



τ-IBR Example: Azithromycin for Prevention of Chronic Obstructive Pulmonary Disease (COPD) Exacerbations Trial

- Fitted Models:
- Logistic regression model for B_i:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -0.681 + 0.558\text{trt} + 0.130\text{age\_10\_new} + 0.497\text{gender} + 0.113\text{fev\_10\_new} + 0.076\text{nowsmk+site variable results}$$

- Beta regression model for Y_i:

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = -0.332 + 0.038\text{trt} + 0.116\text{age\_10\_new} + 0.049\text{gender} + 0.033\text{fev\_10\_new} + 0.001\text{nowsmk+site variable results}$$

Manuscript Presentation of Results (Beta Regression Model for Y_i)

TABLE 3 | Azithromycin for Prevention of COPD Exacerbations Trial: Estimated 1-year-IBR and 1-year-RMST multivariable model parameters with 95% confidence intervals and *p*-values. All models are additionally adjusted for the study site (data not shown).

	Azithromycin (vs. placebo)	Age (per 10 years)	Male (vs. female)	FEV_1 (per 10% predicted)	Current smoker (vs. Ex)
τ-IBR model (MI) (beta regression)					
Fold change[a]	1.022	1.068	1.029	1.020	1.001
95% CI	(0.923, 1.122)	(1.006, 1.131)	(0.928, 1.129)	(0.986, 1.053)	(0.867, 1.134)
p-value	0.655	0.031	0.570	0.249	0.993
τ-IBR model (MI) (logistic regression)					
Odds ratio[b]	1.748	1.139	1.644	1.119	1.079
95% CI	(1.342, 2.276)	(0.966, 1.342)	(1.245, 2.170)	(1.027, 1.220)	(0.754, 1.544)
p-value	<0.001	0.123	<0.001	0.010	0.678

aAmong those experiencing an exacerbation during the 1 year of follow-up, fold change is the ratio of estimated exacerbation-free time during the year when comparing those with versus without a one unit increase in the predictor, assuming all other predictors are zero. Age is centered at 65 years, and the percent of predicted FEV_1 is centered at 40% to aid in interpreting fold changes.
bOdds ratio for remaining exacerbation-free at one year comparing those with versus without a one unit increase in the predictor shown, adjusted for other covariates in the model including treatment group, age, gender, percent of predicted FEV_1, smoking status, and study site.
cPercentage increase in 1-year-RMST for each unit increase of the predictor, adjusted for other covariates in the model.

- Fold change for trt going from z =0 to 1 from MI beta regression model:

$$\frac{e^{\alpha_k}\left(1+e^{\alpha_0+z\alpha_k}\right)}{\left(1+e^{\alpha_0+\alpha_k+z\alpha_k}\right)} = \frac{e^{0.038}\left(1+e^{-0.332}\right)}{\left(1+e^{-0.332+0.038}\right)} = 1.022$$

Amongst those with exacerbations during the year, azithromycin did not significantly increase the percentage of exacerbation-free time during that period (MI: fold change of 1.022 longer exacerbation-free time comparing azithromycin to placebo and adjusted for other factors in the model, 95% CI: 0.923–1.122, p = 0.655).

Manuscript Presentation of Results (Logistic Regression Model for B_i)

TABLE 3 | Azithromycin for Prevention of COPD Exacerbations Trial: Estimated 1-year-IBR and 1-year-RMST multivariable model parameters with 95% confidence intervals and *p*-values. All models are additionally adjusted for the study site (data not shown).

	Azithromycin (vs. placebo)	Age (per 10 years)	Male (vs. female)	FEV_1 (per 10% predicted)	Current smoker (vs. Ex)
τ-IBR model (MI) (beta regression)					
Fold change[a]	1.022	1.068	1.029	1.020	1.001
95% CI	(0.923, 1.122)	(1.006, 1.131)	(0.928, 1.129)	(0.986, 1.053)	(0.867, 1.134)
p-value	0.655	0.031	0.570	0.249	0.993
τ-IBR model (MI) (logistic regression)					
Odds ratio[b]	1.748	1.139	1.644	1.119	1.079
95% CI	(1.342, 2.276)	(0.966, 1.342)	(1.245, 2.170)	(1.027, 1.220)	(0.754, 1.544)
p-value	<0.001	0.123	<0.001	0.010	0.678

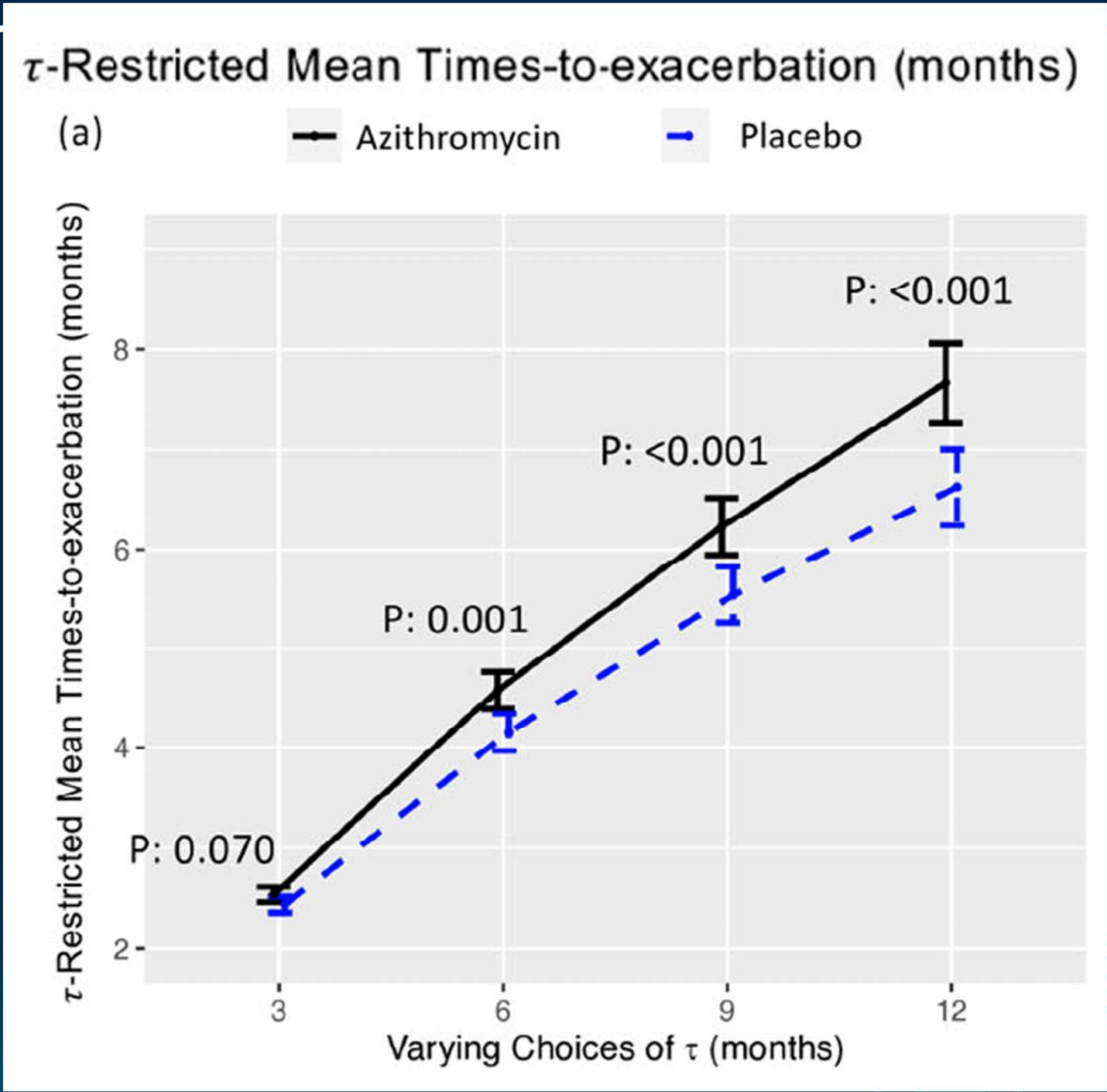
aAmong those experiencing an exacerbation during the 1 year of follow-up, fold change is the ratio of estimated exacerbation-free time during the year when comparing those with versus without a one unit increase in the predictor, assuming all other predictors are zero. Age is centered at 65 years, and the percent of predicted FEV_1 is centered at 40% to aid in interpreting fold changes.
bOdds ratio for remaining exacerbation-free at one year comparing those with versus without a one unit increase in the predictor shown, adjusted for other covariates in the model including treatment group, age, gender, percent of predicted FEV_1, smoking status, and study site.
cPercentage increase in 1-year-RMST for each unit increase of the predictor, adjusted for other covariates in the model.

- Odds ratio for trt going from z =0 to 1 from MI logistic regression model:

$$e^{\beta_k} = e^{0.558} = 1.748$$

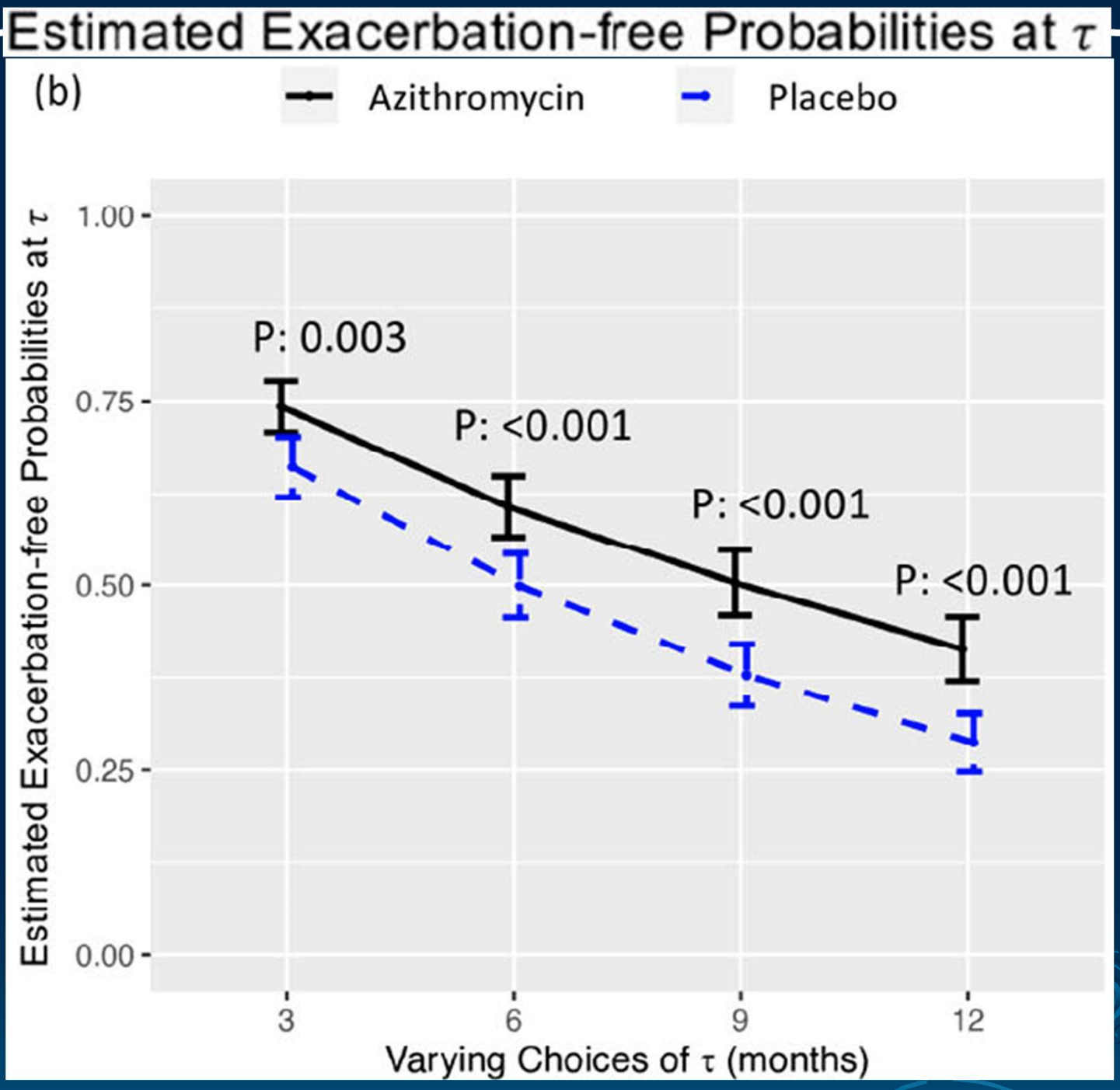
The odds of remaining exacerbation-free at 1 year were significantly higher when taking azithromycin versus placebo, adjusted for other factors in the model (1.748, 95% CI: 1.342–2.276, p < 0.001).

Adjusted Azithromycin Effect on τ −RMST in months (Varying τ)



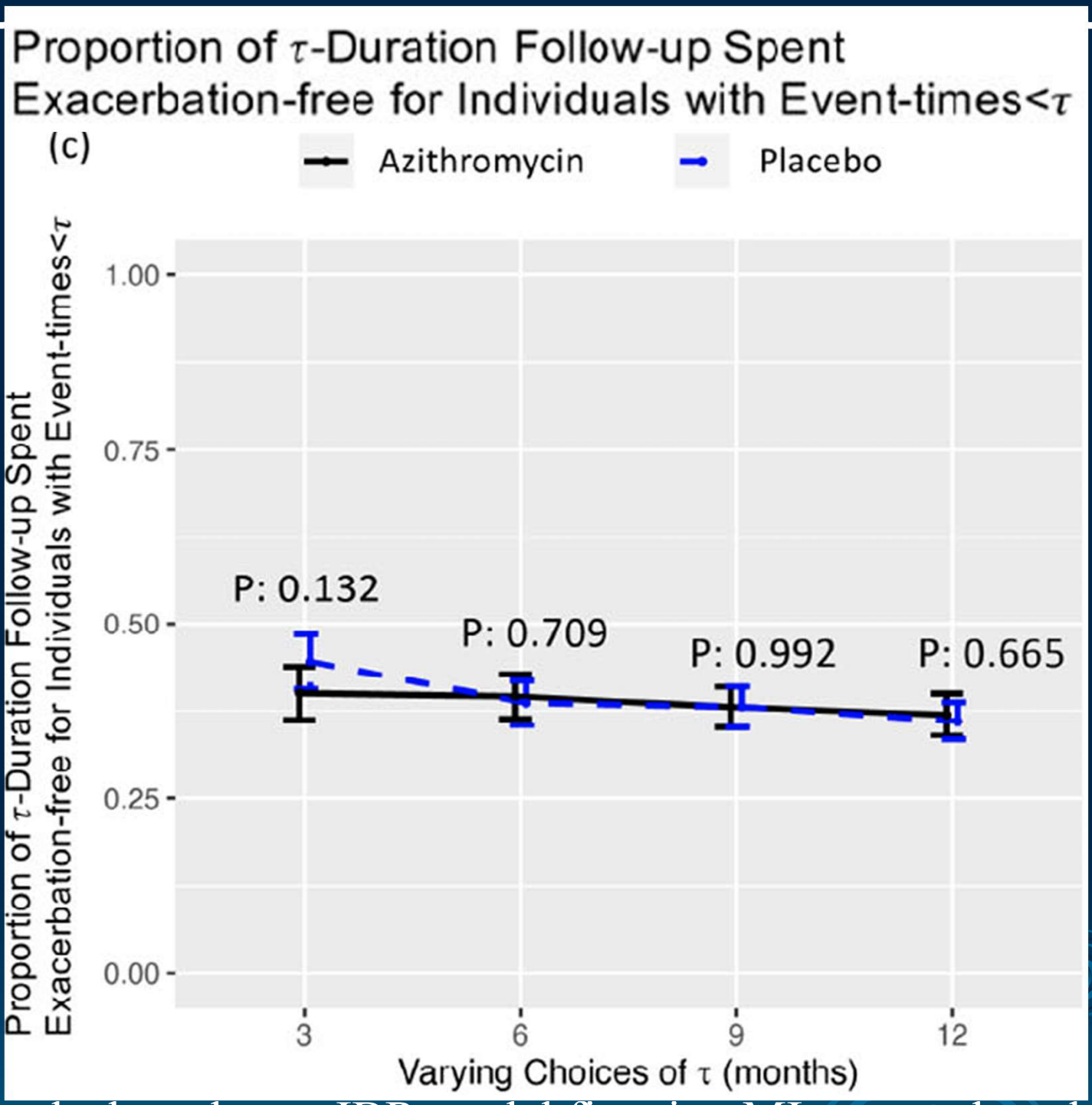
Results based on τ-IBR model fit using MI approach and assuming average cohort values for confounders: age 65 years, 59% probability of male, 40% FEV1 % predicted, 22% chance of being a current smoker and study site probabilities (not shown).

Adjusted Azithromycin Effect on Exacerbation-free Probabilities at τ (Varying τ)



Results based on τ-IBR model fit using MI approach and assuming average cohort values for confounders: age 65 years, 59% probability of male, 40% FEV1 % predicted, 22% chance of being a current smoker and study site probabilities (not shown).

Adjusted Azithromycin Effect on Proportion of τ-Duration Follow-up Spent Exacerbation-free for Individuals with Event-times $< \tau$ (Varying τ)



Results based on τ-IBR model fit using MI approach and assuming average cohort values for confounders: age 65 years, 59% probability of male, 40% FEV1 % predicted, 22% chance of being a current smoker and study site probabilities (not shown).

How to Use Github Site for Ann Arbor ASA Meeting 1-28-2025

- Go to Ann Arbor ASA Meeting Github Site: https://github.com/yezow/tau-IBR_ASA
- Create a folder with the following :
  - Dataset: ex_data_example.csv
  - Paper1_func.R
  - MI_function.R
  - Your current R code file
  - Following code should work with R Studio

Load a Few R Packages

```
library(foreach)
library(doParallel)
library(parallel)
library(boot)
library(purrr)
library(survival)
library(survminer)
library(data.table)
library(survRM2)
library(coda)
library(numDeriv)
library(matrixStats)
library(betareg)
library(simcausal)
library(latex2exp)
library(dplyr)
library(base)
library(eoffice)
library(ggplotify)
library(gridExtra)
library(cowplot)
library(pheatmap)
library(viridis)
library(RColorBrewer)
library(ggplot2)
library(tikzDevice)
library(ggpubr)
library(grid)
```

Data Preparation

```
options(warn=0) #keep running the code with warnings
current_path = rstudioapi::getActiveDocumentContext()$path
setwd(dirname(current_path ))

#X and Delta variables must use these variable names

#Need to define all variables in mu model
Z_mu <- data.matrix(dplyr::select(Data_cens,
             trt,
            age_10_new,
            gender,
            fev1_10_new,
            nowsmk,
            latitude_38,latitude_39,latitude_40,
            latitude_42,latitude_44))
Z_mu_length=ncol(Z_mu)

# Need to define all variables in pi model (same for this
example, but need not be)

Z_pi <- Z_mu
Z_pi_length=ncol(Z_pi)

nsubj<-nrow(Z_mu)

#rows that need multiple imputation

row_MI <- which(Data_cens$X<tau &
                      Data_cens$Delta==0)
```

Fitting τ-IBR model

```
registerDoParallel(cores=4)
import::from("numDeriv", "hessian")

#set initial model parameters for estimation (mu model then pi model)
initial_para_est <-rep(0, Z_mu_length+ Z_pi_length)

#run code for τ-IBR model once all variables on previous slide defined
source("Paper1_func.R")
source("MI_function.R")

#MI coefficient estimates (mu then pi model)
coef_MI<-MI_result[[1]]

#corresponding MI coefficient variances
var_MI<-MI_result[[2]]

#EM coefficient estimates (mu then pi model)
EM_result<-MI_impute_converge()
coef_EM <-EM_result[[1]]

#corresponding EM coefficient variances
var_EM <- EM_var(w,coef_EM)
```

MI Model Results

```
results_pi=round(rbind(coef_pi=coef_MI[(ncol(Z_mu)+2):
(ncol(Z_mu)+ncol(Z_pi)+2)],var_pi=var_MI[(ncol(Z_mu)+
2):(ncol(Z_mu)+ncol(Z_pi)+2)]),digits=3)
colnames(results_pi)=colnames(results_mu)
results_pi
```

```
> results_pi
        intercept   trt age_10_new gender fev1_10_new nowsmk latitude_38 latitude_39 latitude_40
coef_pi    -0.681 0.558       0.13  0.497       0.113  0.076      -0.525      -1.019      -0.474
var_pi     -0.003 0.001       0.00  0.001       0.000 -0.001       0.001       0.003       0.002
        latitude_42 latitude_44
coef_pi      -0.870      -0.777
var_pi        0.002       0.003
```

```
results_mu=round(rbind(coef_mu=coef_MI[1:(ncol(Z_mu)
+1)],var_mu=var_MI[1:(ncol(Z_mu)+1)]),digits=3)
colnames(results_mu)<-
c("intercept","trt","age_10_new","gender","fev1_10_new",
"nowsmk","latitude_38","latitude_39","latitude_40","latitud
e_42","latitude_44")
results_mu
```

```
        intercept    trt age_10_new gender fev1_10_new nowsmk latitude_38 latitude_39 latitude_40
coef_mu    -0.332  0.038       0.116  0.049       0.033  0.001      -0.195      -0.187      -0.270
var_mu      0.019 -0.004       0.000 -0.005       0.000 -0.004      -0.013      -0.012      -0.014
        latitude_42 latitude_44
coef_mu      -0.409      -0.577
var_mu       -0.014      -0.014
```

Example of Fold Change Code for Trt from Beta Regression Model for $Y_i = \frac{T_i}{\tau}$ given $T_i < \tau$

```
#function called fold_change helps with calculations for MI estimates. E.g. for trt:

fold_change_trt <- fold_change(coef_MI[1:2])

#Corresponding variance
   var_fold_change_trt <-
fold_change_var(coef_MI[1:2],var_MI[1:2,1:2])

#Lower and Upper 95% CI
   fold_change_CI_lower<-c(fold_change_trt-
1.96*sqrt(var_fold_change_trt))
   fold_change_CI_upper<-
c(fold_change_trt+1.96*sqrt(var_fold_change_trt))

#p-value
P_mu<-2*(1-
pnorm(abs(coef_MI[2]),mean=0,sd=sqrt(var_MI[2,2])))

#Display results
beta_model_results_trt<-rbind(fold_change_trt,
                      fold_change_CI_lower,
                      fold_change_CI_upper,
                      P_mu)
round(beta_model_results_trt, digits=3)
```

Change highlighted 2's to covariate position in Z_mu to get results for remaining variables

```
fold_change_trt       1.022
fold_change_CI_lower  0.923
fold_change_CI_upper  1.122
P_mu                  0.655
```

Odds Ratio Results From Logistic Regression Model for $B_i = I(T_i \geq \tau)$

```
pi_model_results<-function(coef,var){
   odds_ratio_est<-exp(coef[13:17])
   odds_ratio_lower <- numeric()
   odds_ratio_upper <- numeric()
   for (j in 13:17){
     odds_ratio_lower <-
append(odds_ratio_lower,exp(coef[j]-1.96*sqrt(var[j,j])))
     odds_ratio_upper <-
append(odds_ratio_upper,exp(coef[j]+1.96*sqrt(var[j,j])))
   }
   P_pi<-numeric()
   for (i in c(13:17)){
     P_pi<-append(P_pi,(2*(1-
pnorm(abs(coef[i]),mean=0,sd=sqrt(var[i,i])))))
   }
   pi_model_result<-rbind(odds_ratio_est,
                          odds_ratio_lower,
                          odds_ratio_upper,
                           P_pi)
   return(pi_model_result)
}

pi_model_results_MI<-pi_model_results(coef_MI,var_MI)
```

13:17 th elements in overall covariate matrix

	trt	age_10_new	gender	fev1_10_new	nowsmk
odds_ratio_est	1.747878e+00	1.1386911	1.6437379415	1.11935771	1.0788680
odds_ratio_lower	1.342428e+00	0.9656047	1.2448368429	1.02702492	0.7540686
odds_ratio_upper	2.275784e+00	1.3428034	2.1704646965	1.21999151	1.5435679
P_pi	3.369541e-05	0.1226049	0.0004579197	0.01025489	0.6778524

Summary

- The τ -IBR model should be considered if modeling a cohort of participants with many $\min(\tau, T_i)$ events reaching τ.
- Our simulation results indicate that traditional τ –RMST models give very noisy results if important covariates predict those who are event-free at τ.
- Seeing regression results for the two different components of the τ -IBR model can be helpful in understanding treatment effects.