IPS / Translational Bioinformatics

Composite Biomarker Analysis of Patients with Urothelial Carcinoma

April 10, 2024

Abraham Apfel
Translational Bioinformatics / Bioinformatics Methodology

Bristol Myers Squibb™

Key Research Question

- Estimating associations between 13 biomarkers and Disease-Free Survival (DFS) in Urothelial Carcinoma (UC) patients treated with adjuvant therapy

  - Do the estimated associations depend on whether:
    - the patient received Treatment or Placebo?
    - we adjust for other biomarkers?

Acknowledgements

Scott Chasalow

Jun Li

Justin David

Bhakti Dwivedi

Our Cohort

- Phase 3 Randomized Clinical Trial of 699 patients
  - 2 treatment arms
- Whole Exome Sequencing (WES)
  - N = 458
  - Used to derive TMB
  - Identified genes with most mutations
- RNA-Seq
  - N = 323
  - Genome-wide single gene and gene signature analyses
- **Composite Analysis**
  - **Used literature and stratification variables to select a subset of 13 biomarkers for which to take deeper dive and run composite analysis.**

Statistical Considerations (1)

Challenge	Solution
Straight-line assumption can yield misleading results when assessing interactions among variables[1]	**Transform each continuous variable into 2 basis functions to allow for cubic splines in model**
Previous studies imply different associations between PDL1_IHC (2 measures) = 0 and PDL1_IHC > 0. Not sufficiently accounted for by splines	Force model to allow for point of discontinuity at PDL1_IHC = 0 by including extra terms in model
Train model to focus on discovering interactions between variables and treatment	Parameterize treatment as -1/+1 instead of standard 0/1 “dummy” coding[2]

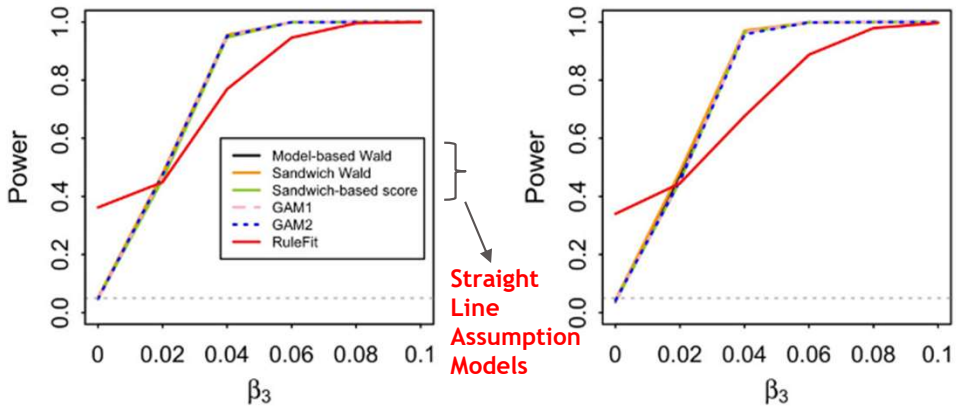
Simulations for Continuous Outcome:

Straight Line Assumption Leads to High False Discovery Rate in Many Non-Linear Scenarios

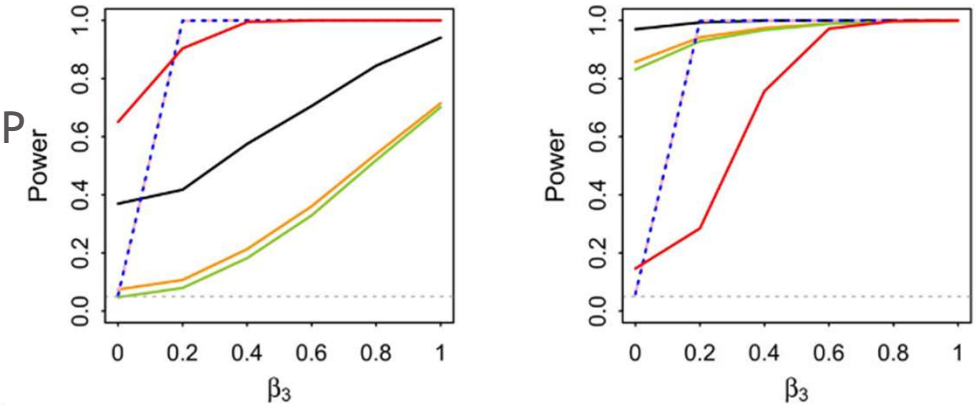
X_1 and X_2 Independent

X_1 and X_2 Correlated

TRUE RELATIONSHIP
Linear



TRUE RELATIONSHIP
Quadratic



- N = 500
- β_3 is coefficient for X1*X2, the interaction term
- "RuleFit" is method proposed by Friedman, et al for discovering interactions
- Intercept (Power for β_3 = 0) is Type I error

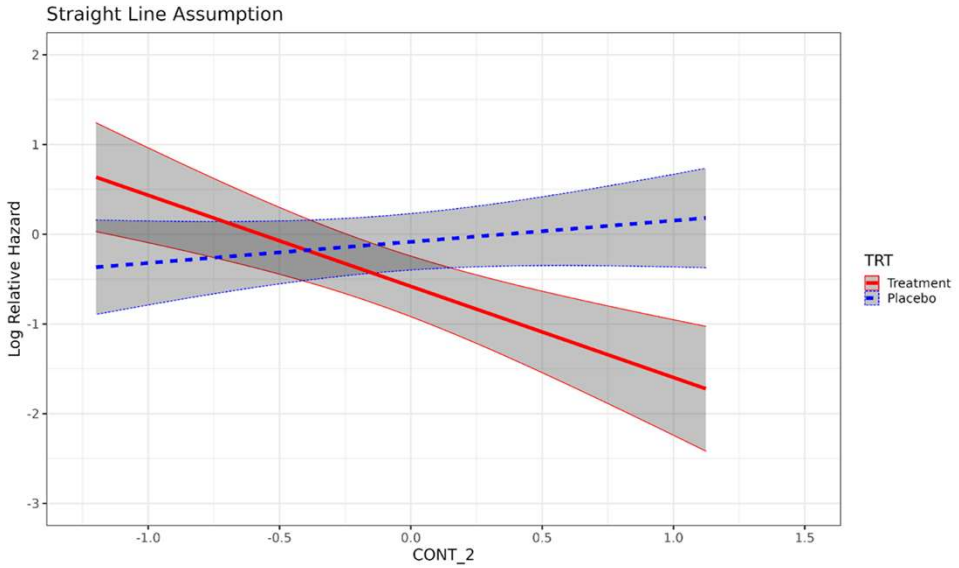
- NOTE: Results were similar for Log and Exponential relationship with outcome
- For Binary outcome, even when predictors were independent there is high Type I Error

1. Zhang, M., Yu, Y., Wang, S., Salvatore, M., G. Fritsche, L., He, Z., & Mukherjee, B. (2020). Interaction analysis under misspecification of main effects: Some common mistakes and simple solutions. *Statistics in Medicine*, *39*(11), 1675-1694.

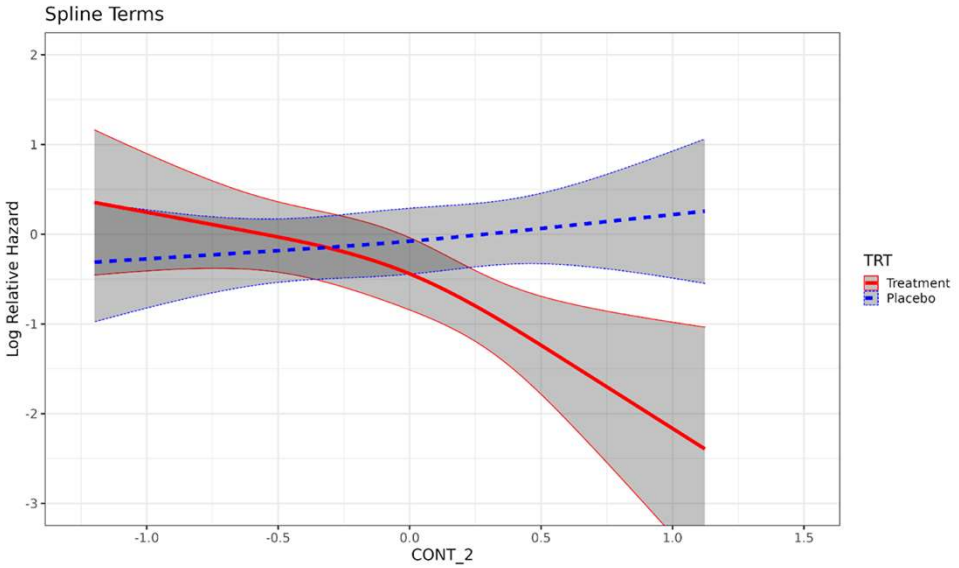
Model with One Continuous Biomarker at a Time

"Simple" Model: Surv(DFS, Censor) ~ Biomarker*TRT + Stratification_Variables

Straight Line Assumption



Restricted Cubic Spline



Statistical Considerations (1)

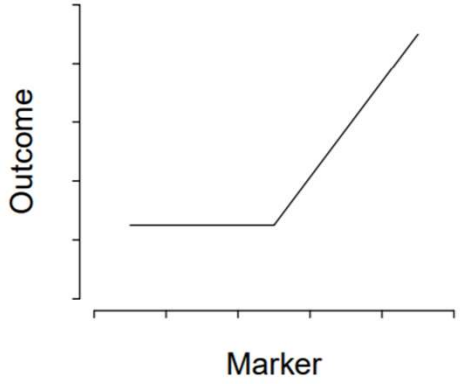
Challenge	Solution
Straight-line assumption can yield misleading results when assessing interactions among variables[1]	Transform each continuous variable into 2 basis functions to allow for cubic splines in model
Previous studies imply different associations between PDL1_IHC (2 measures) = 0 and PDL1_IHC > 0. Not sufficiently accounted for by splines	**Force model to allow for point of discontinuity at PDL1_IHC = 0 by including extra terms in model**
Train model to focus on discovering interactions between variables and treatment	Parameterize treatment as -1/+1 instead of standard 0/1 “dummy” coding[2]

Implications of Dichotomizing a Continuous Variable

Continuous (Can be modeled via Splines)

Can Occur in Biology

Not Handled by Dichotomization



Dichotomized

Unlikely to Occur

Assumed in Much of Biomarker Research

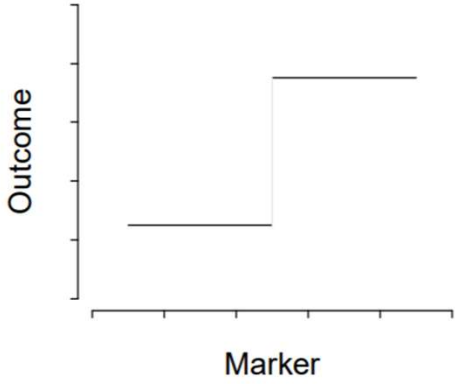


Figure 18.2: Two kinds of thresholds. The pattern on the left represents a discontinuity in the first derivative (slope) of the function relating a marker to outcome. On the right there is a lowest-order discontinuity.

- Prior studies suggest that each PDL1 measure has point of discontinuity at 0
- Still treat as continuous (i.e., do not assume a flat line) for values greater than 0

Figure copied from:

http://hbiostat.org/doc/bbr.pdf (Frank Harrell)

Statistical Considerations (1)

Challenge	Solution
Straight-line assumption can yield misleading results when assessing interactions among variables[1]	Transform each continuous variable into 2 basis functions to allow for cubic splines in model
Previous studies imply different associations between PDL1_IHC (2 measures) = 0 and PDL1_IHC > 0. Not sufficiently accounted for by splines	Force model to allow for point of discontinuity at PDL1_IHC = 0 by including extra terms in model
Train model to focus on discovering interactions between variables and treatment	**Parameterize treatment as -1/+1 instead of standard 0/1 "dummy" coding[2]**

Statistical Considerations (2)

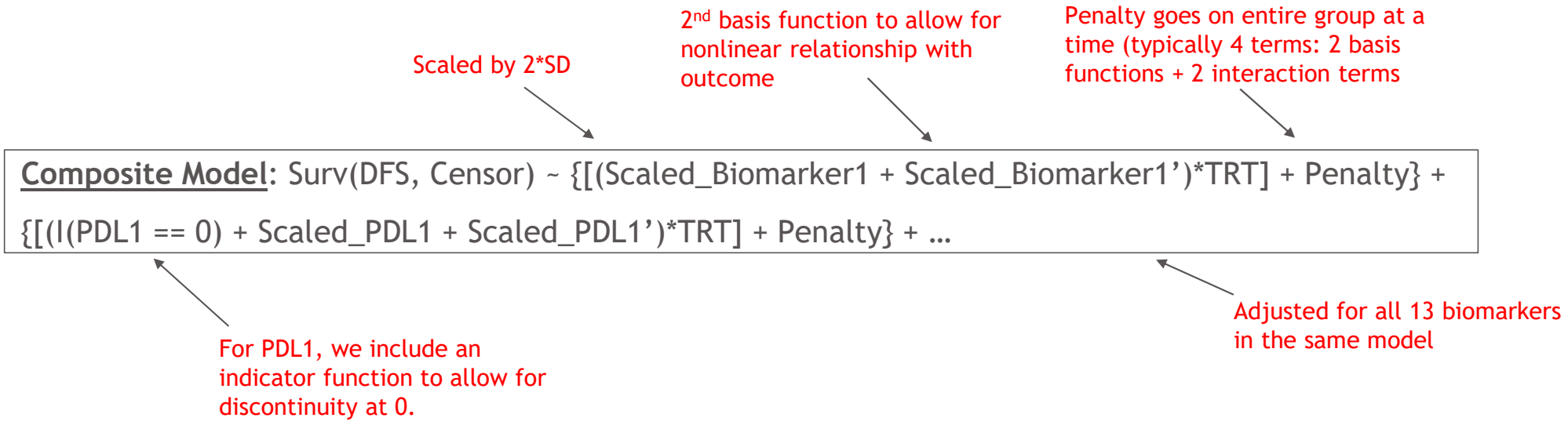
Challenge	Solution
Small sample size (N = 323) relative to number of predictors (13 + Treatment) and terms (2 basis functions and interaction for each continuous predictor)	Penalized Regression
Maintain hierarchy while allowing for multiple basis functions (for splines) and interaction with treatment in penalized setting	Use Group Lasso[3]: Each group consists of all terms associated with a given variable
Different scaling of categorical and continuous predictors makes it difficult to infer relative importance of variables in penalized setting	Scale each continuous variable by dividing by 2*Standard Deviation instead of typical scaling of 1*Standard Deviation[4]
Appropriate visualizations to assess effect size when allowing for non-straight-line effects	Make partial effects plots and coefficient trajectory plots

Intuition Behind Scaling by 2*SD_4

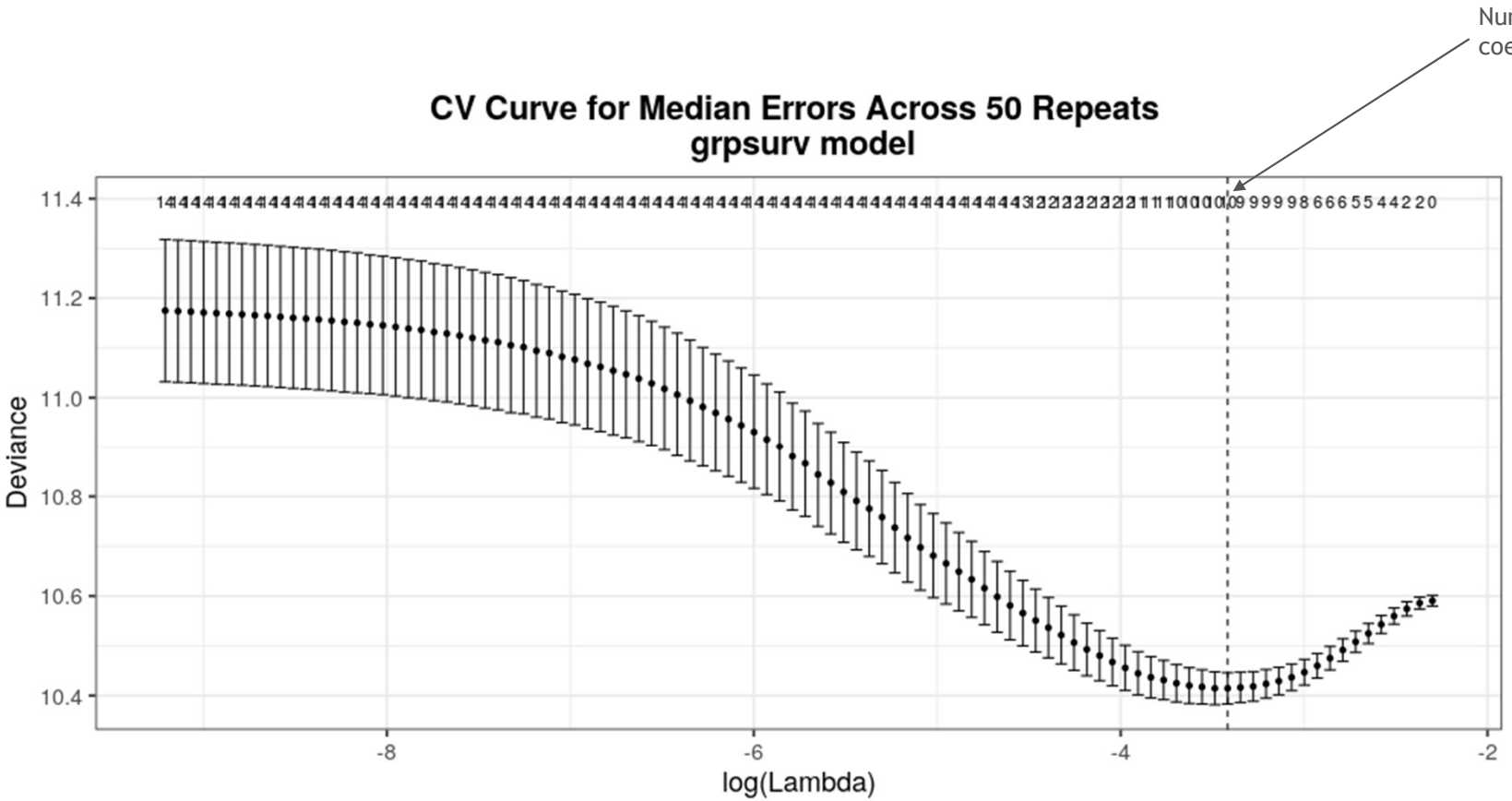
- Each continuous variable will have SD = 0.5
- Binary variable, X, with P(X = 1) = 0.5 has SD = 0.5
  - Thus by dividing continuous predictors by 2*SD, they will be on same scale as categorical predictor with P(X = 1) = 0.5
- Even categorical predictor with skewed distribution, e.g. Pr(X = 1) = 0.3, has SD = 0.45
  - Thus, unless categorical predictor has "extremely skewed" distribution, dividing by 2*SD is sufficient to put continuous variables on "similar" scale
    - In our dataset, most skewed variable had Pr(X = 1) = 0.21, with SD = 0.4

4. Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, *27*(15), 2865-2873.

Composite Model



Selecting lambda (degree of penalization)



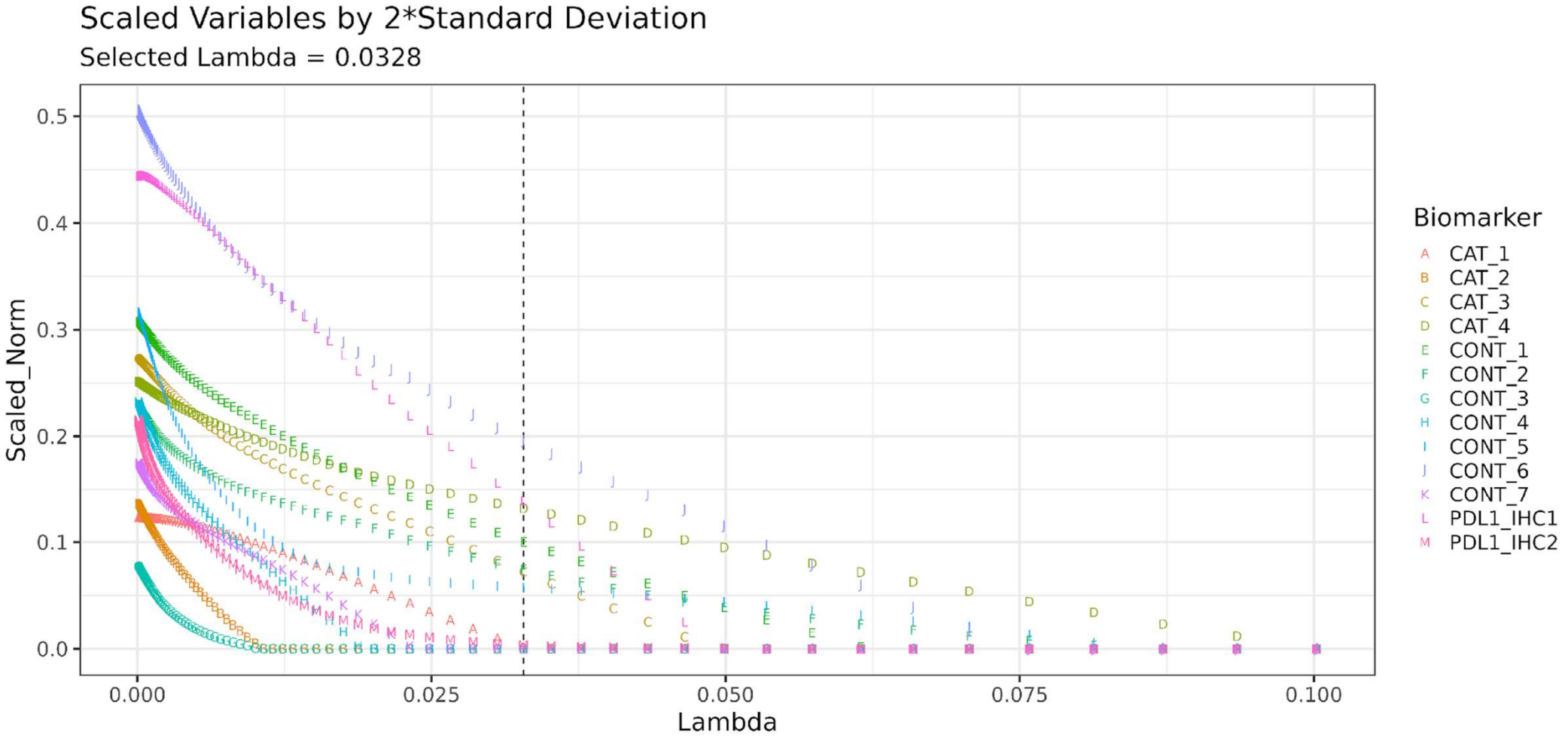
We chose the Lambda with minimum Deviance:

Lambda = 0.0328 (log(lambda) = -3.42)

Statistical Considerations (2)

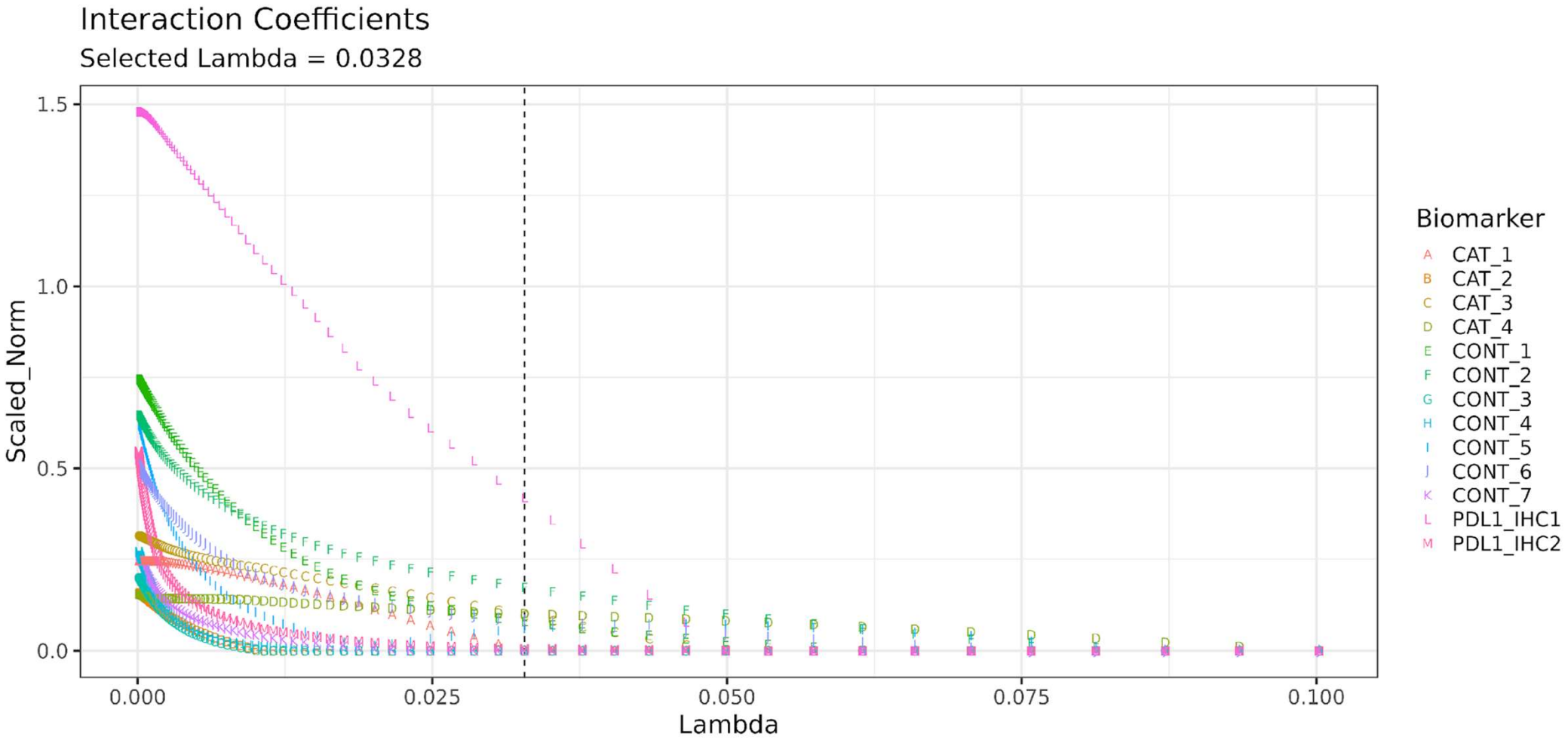
Challenge	Solution
Small sample size (N = 323) relative to number of predictors (13 + Treatment) and terms (2 basis functions and interaction for each continuous predictor)	Penalized Regression
Maintain hierarchy while allowing for multiple basis functions (for splines) and interaction with treatment in penalized setting	Use Group Lasso[3]: Each group consists of all terms associated with any given variable
Different scaling of categorical and continuous predictors makes it difficult to infer relative importance of variables in penalized setting	Scale each continuous variable by dividing by 2*Standard Deviation instead of typical scaling of 1*Standard Deviation[4]
Appropriate visualizations to assess effect size when allowing for non-straight-line effects	**Make partial effects plots and coefficient trajectory plots**

Trajectories of Group Coefficient Norms



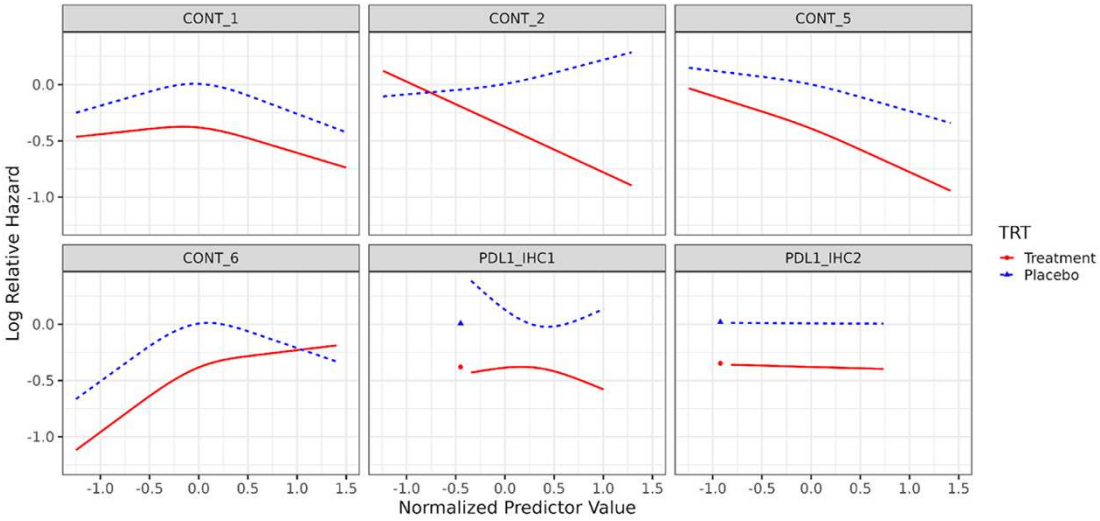
Scaled_Norm = $\sqrt{(\sum B\text{^}2)/k}$, where B = Coefficient and k = Number of coefficients for a given predictor

Trajectories of Interaction Coefficient Norms

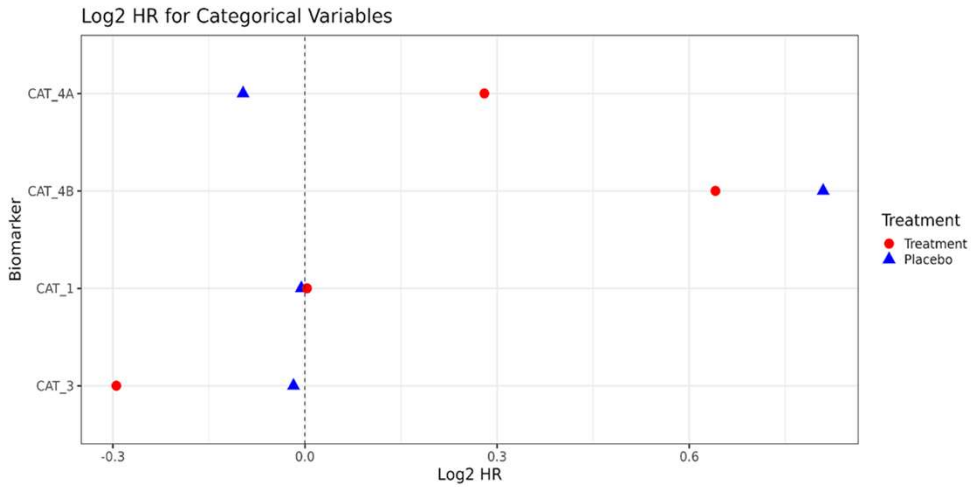


Results from Model: Partial Effects Plots

Continuous Variables

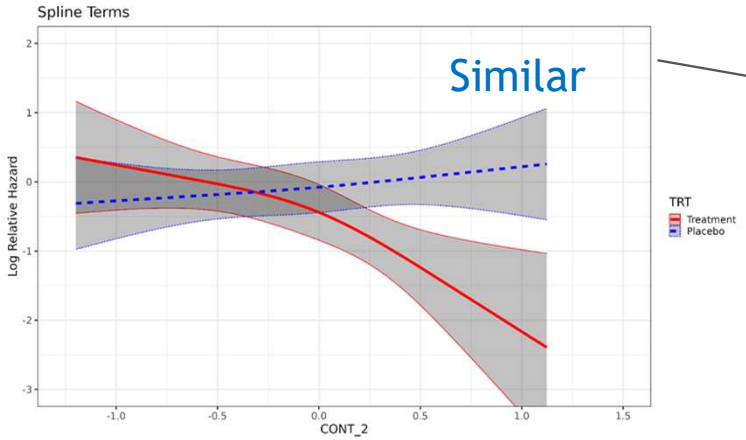


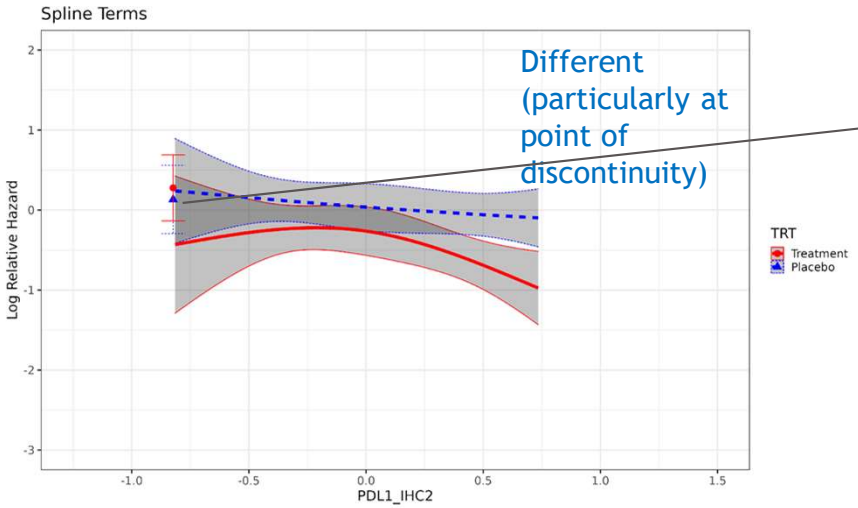
Categorical Variables



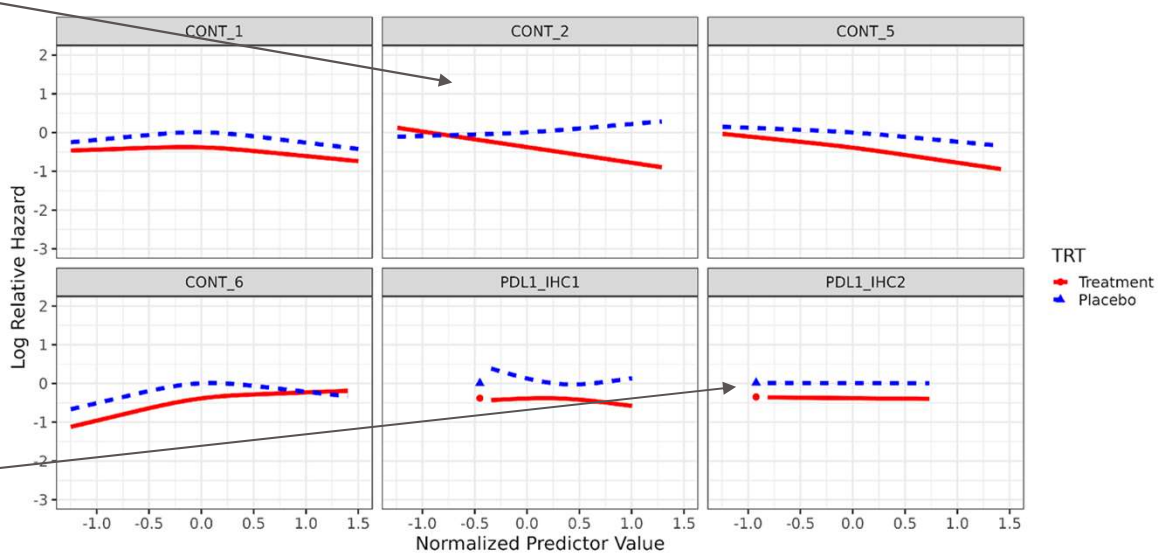
Composite Model More Impactful for Some Biomarkers Than Others

One-At-A-Time Model



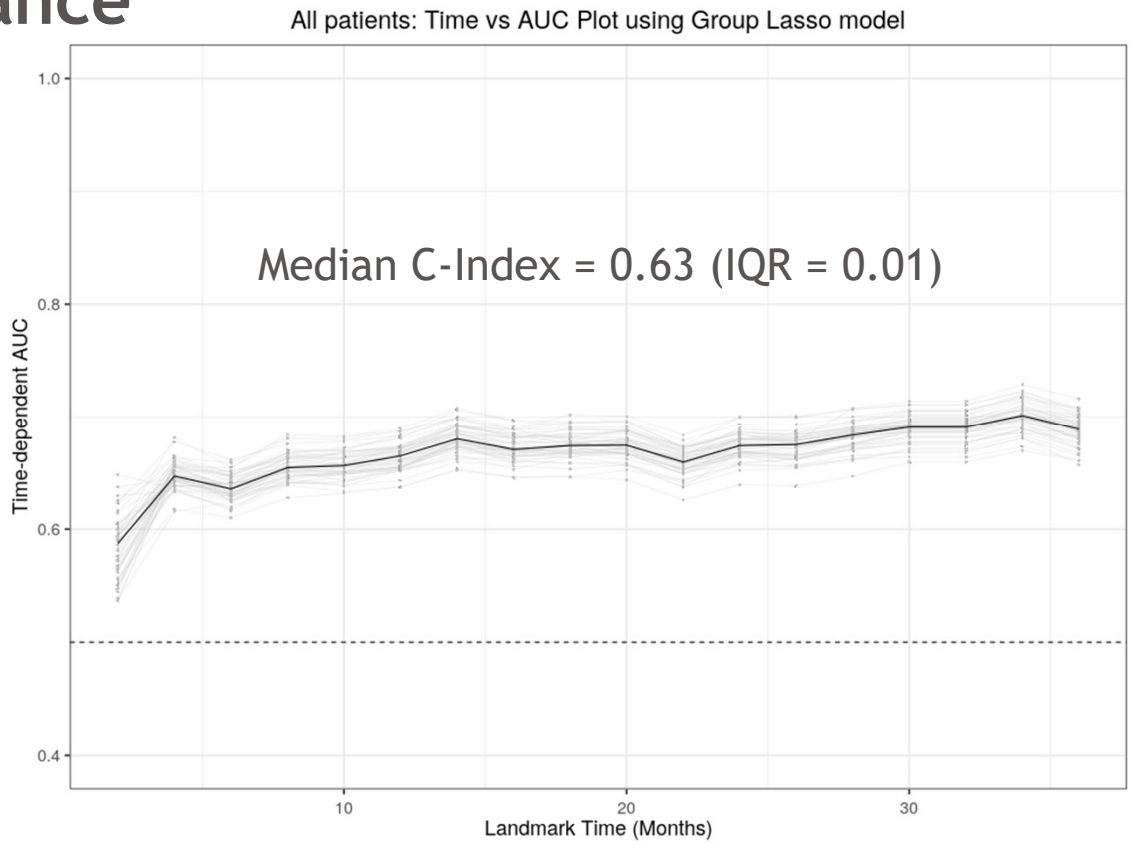


Composite Model



Model Performance

Note: Performance Metrics are from 50 repeats of 10-fold CV

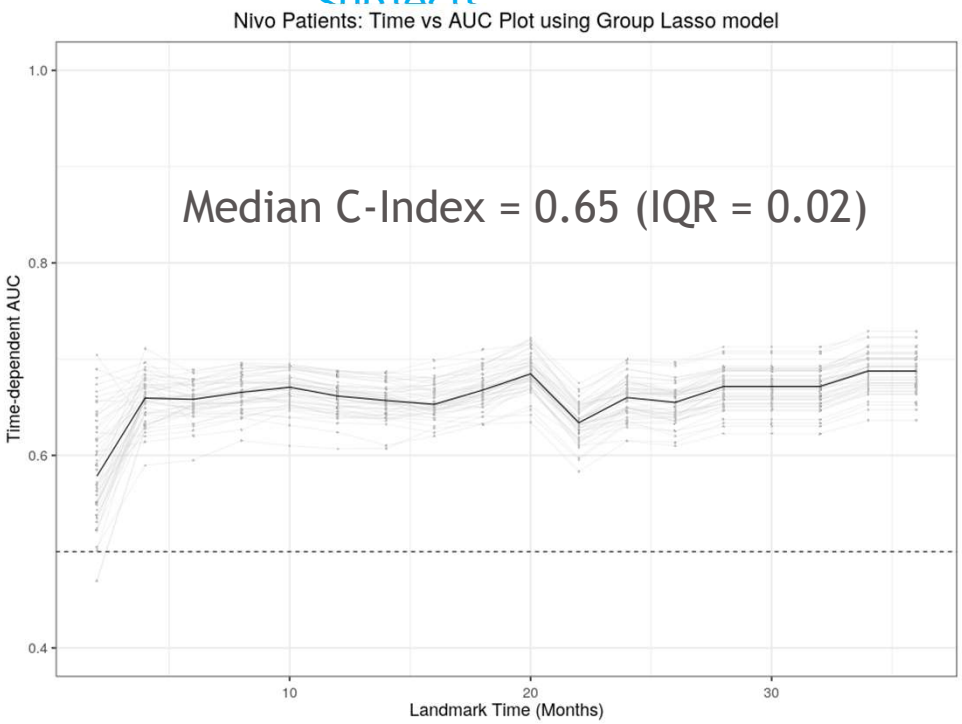


Distribution of DFS

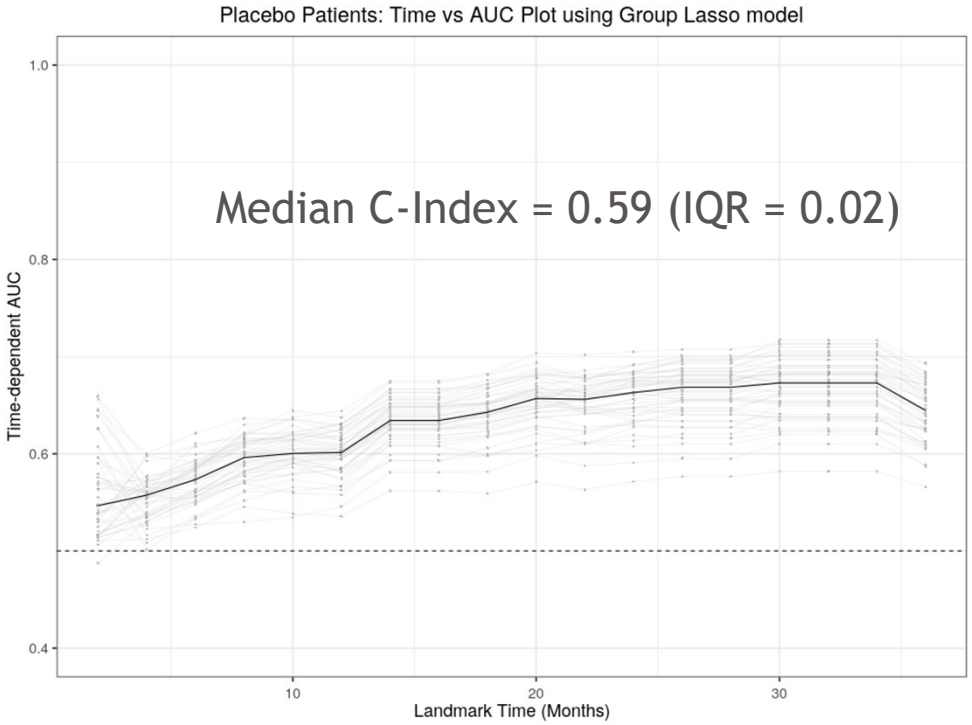
10%	25%	50%	75%	90%
2.6	3.4	13.7	25.8	36.9

Model Performance, by Arm

Treatment Subjects



Placebo Subjects



Note: Performance Metrics are from 50 repeats of 10-fold CV

Statistical Considerations (1)

Challenge	Solution
Straight-line assumption can yield misleading results when assessing interactions among variables[1]	Transform each continuous variable into 2 basis functions to allow for cubic splines in model
Previous studies imply different associations between PDL1_IHC (2 measures) = 0 and PDL1_IHC > 0. Not sufficiently accounted for by splines	Force model to allow for point of discontinuity at PDL1_IHC = 0 by including extra terms in model
Train model to focus on discovering interactions between variables and treatment	Parameterize treatment as -1/+1 instead of standard 0/1 “dummy” coding[2]

Statistical Considerations (2)

Challenge	Solution
Small sample size (N = 323) relative to number of predictors (13 + Treatment) and terms (2 basis functions and interaction for each continuous predictor)	Penalized Regression
Maintain hierarchy while allowing for multiple basis functions (for splines) and interaction with treatment in penalized setting	Use Group Lasso[3]: Each group consists of all terms associated with any given variable
Different scaling of categorical and continuous predictors makes it difficult to infer relative importance of variables in penalized setting	Scale each continuous variable by dividing by 2*Standard Deviation instead of typical scaling of 1*Standard Deviation[4]
Appropriate visualizations to assess effect size when allowing for non-straight-line effects	Make partial effects plots and coefficient trajectory plots

Conclusions: Clinical

- In our composite model:
  - The effect of PDL1_IHC1 and CONT_2 on DFS differed greatly between the Treatment and Placebo arms
  - CONT_1, CONT_5, CONT_6, and CAT_4 were possibly prognostic: they had strong associations with DFS but NOT strong evidence that the associations differed between treatment arms

Conclusions: Methods

- The results for some biomarkers changed greatly between the “one continuous biomarker at a time” model and the composite model
  - —Most notably, CONT_3 and PDL1_IHC2 were found to have minimal association with DFS in the composite model but each appeared to have strong association in the one-at-a-time models

- The detailed modeling choices had important influences on our conclusions
  - —Relationship of PDL1_IHC1 and PDL1_IHC2 to DFS changed greatly after allowing for discontinuity at 0
  - —Appropriate scaling changed the relative importance of several variables
  - —The straight-line assumption appeared unreasonable for several variables

References

1. Zhang, M., Yu, Y., Wang, S., Salvatore, M., G. Fritsche, L., He, Z., & Mukherjee, B. (2020). Interaction analysis under misspecification of main effects: Some common mistakes and simple solutions. *Statistics in Medicine*, *39*(11), 1675-1694.
2. Tian, L., Alizadeh, A. A., Gentles, A. J., & Tibshirani, R. (2014). A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association*, *109*(508), 1517-1532.
3. Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *68*(1), 49-67.
4. Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, *27*(15), 2865-2873.
5. Huang J, Breheny P, and Ma S. (2012). A selective review of group selection in high dimensional models. *Statistical Science*, **27**: 481-499. doi: 10.1214/12-sts392

Backup

Methods - 1

- Parameterized Treatment as +1/-1 (instead of standard 0/1 “dummy coding”) to train model to focus on identifying interactions with treatment[2]

- Applied Grouped lasso-regularized Cox PH regression[5]
  - Imputed missing values via single imputation (small amount of missing)
  - Allowed continuous variables to have non-linear relationship with outcome via cubic splines (via “rcs” function from “rms” package)
  - Scaled continuous variables to have mean = 0 and std = 0.5 to achieve similar scaling as categorical variables[4]
  - Each measure for PDL1 was modeled using Ind(x = 0) to allow for a discontinuity between 0 and 1, together with cubic splines for the x > 0 portion of the domain, and interaction with treatment for both parts.
  - Treatment was constrained to have a non-zero coefficient

2 Tian L, Alizadeh AA, Gentles AJ, Tibshirani R. A Simple Method for Estimating Interactions between a Treatment and a Large Number of Covariates. *J Am Stat Assoc*. 2014;109(508):1517-1532. doi:10.1080/01621459.2014.951443

4 Gelman A. Scaling regression inputs by dividing by two standard deviations. Stat Med. 2008 Jul 10;27(15):2865-73. doi: 10.1002/sim.3107. PMID: 17960576.

5 Huang J, Breheny P, and Ma S. (2012). A selective review of group selection in high dimensional models. *Statistical Science*, **27**: 481-499. doi: 10.1214/12-sts392

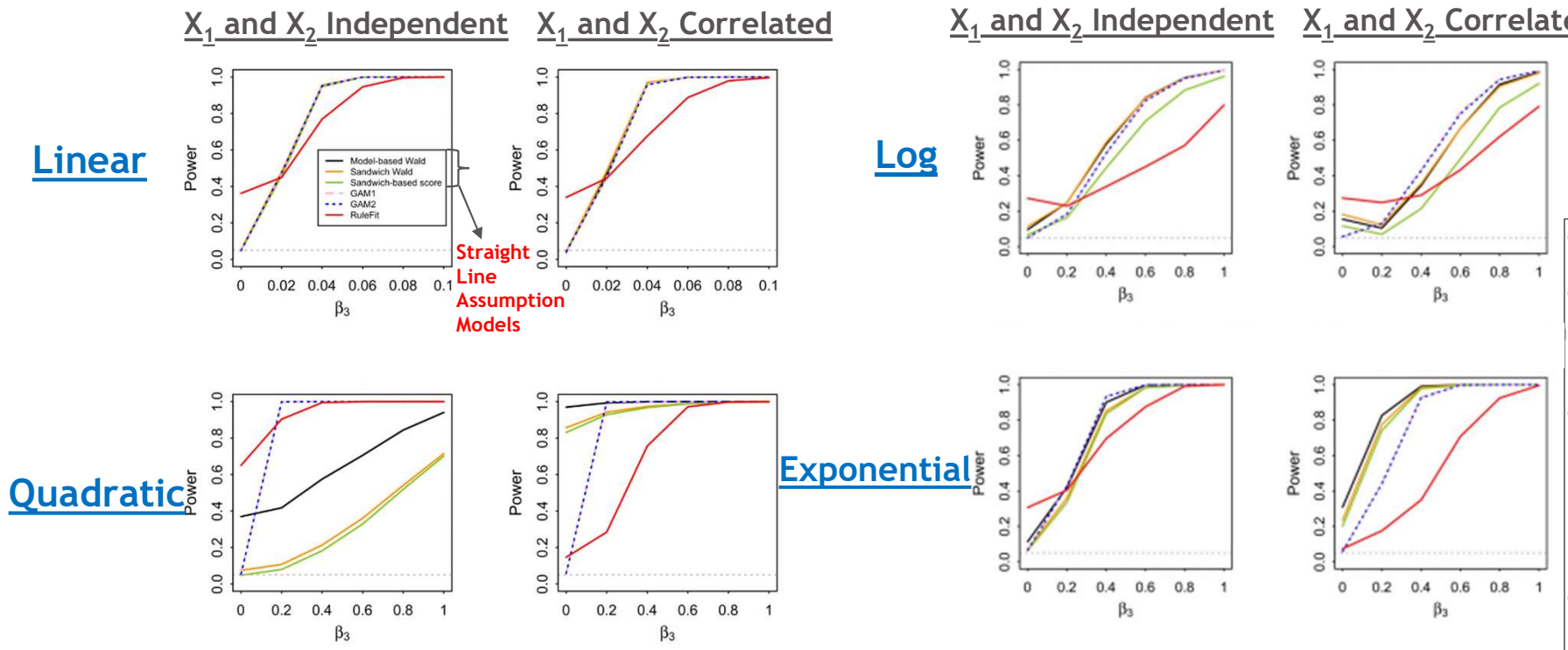
Methods - 2

- Chose the tuning parameter (lambda) value that had minimum median Deviance from 50 repeated iterations of 5-fold cross validation (CV) across the full dataset (inner loop of CV)
- Model Performance was assessed via C-index and plots of AUC (of time-dependent ROC curve) vs Landmark Time, estimated from 50 repeats of 10-fold CV (outer loop)
- Key Packages:
  - grpreg
  - SurvivalROC

Bioinformatics Details for RNA-Seq

- TMM-normalized CPM
- Filtered ~35,000 genes with log2(CPM) <0.5
  - ~24,000 genes remained in analysis
- Gene Set Enrichment Analysis (GSEA)
  - Genes were ranked by logHR (IQR) for expression in placebo arm – logHR (IQR) for expression in treatment arm
  - Looked for patterns of enrichment based on above criterion to identify interesting pathways
  - GSEA used 50 hallmark gene sets from Molecular Signature Data Base (MSigDB)

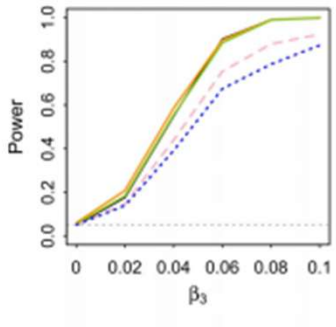
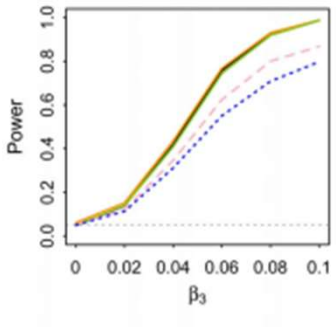
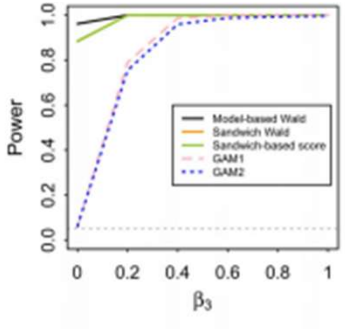
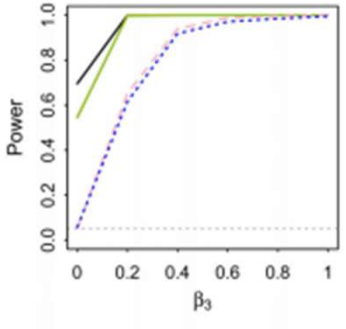
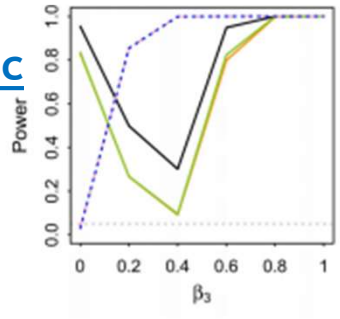
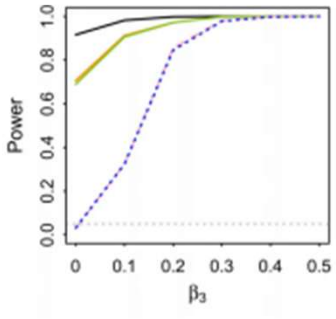
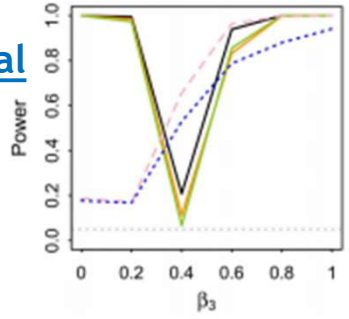
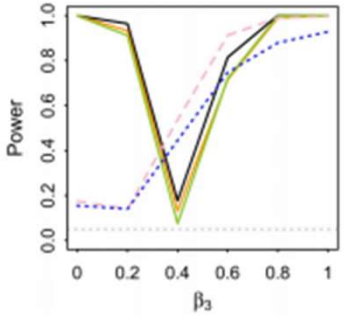
Simulations for Continuous Outcome:
Straight Line Assumption Leads to High False Discovery Rate in Many Non-Linear Scenarios



- N = 500
- β_3 is coefficient for X1*X2, the interaction term
- "RuleFit" is method proposed by Friedman, et al for discovering interactions
- Intercept (Power for $\beta_3 = 0$) is Type I error

1. Zhang, M., Yu, Y., Wang, S., Salvatore, M., G. Fritsche, L., He, Z., & Mukherjee, B. (2020). Interaction analysis under misspecification of main effects: Some common mistakes and simple solutions. *Statistics in Medicine*, *39*(11), 1675-1694.

Simulations for Binary Outcome:
Straight Line Assumption Leads to High False Discovery Rate in ALL Simulated Non-Linear Scenarios

	X_1 and X_2 Independent	X_1 and X_2 Correlated	X_1 and X_2 Independent	X_1 and X_2 Correlated
Linear			Log 	
Quadratic			Exponential 	

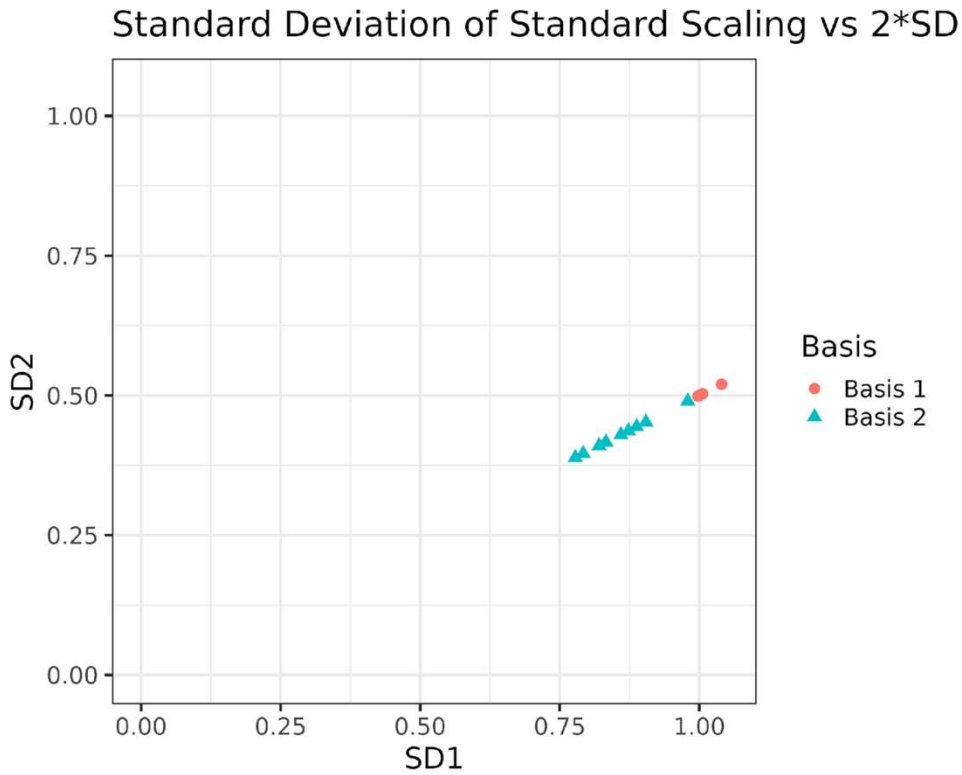
- N = 2000
- β_3 is coefficient for X1*X2, the interaction term
- Intercept (Power for $\beta_3 = 0$) is Type I error

Parameterization of Treatment

- Parameterize treatment as +1/-1 instead of standard dummy coding
- Objective is to estimate $\widehat{\Delta z}$, the causal treatment effect for patients with covariates **z**
  - $\Delta z = E(Y^{(1)} - Y^{(-1)} | \mathbf{Z} = \mathbf{z})$

2. Tian, L., Alizadeh, A. A., Gentles, A. J., & Tibshirani, R. (2014). A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association*, *109*(508), 1517-1532.

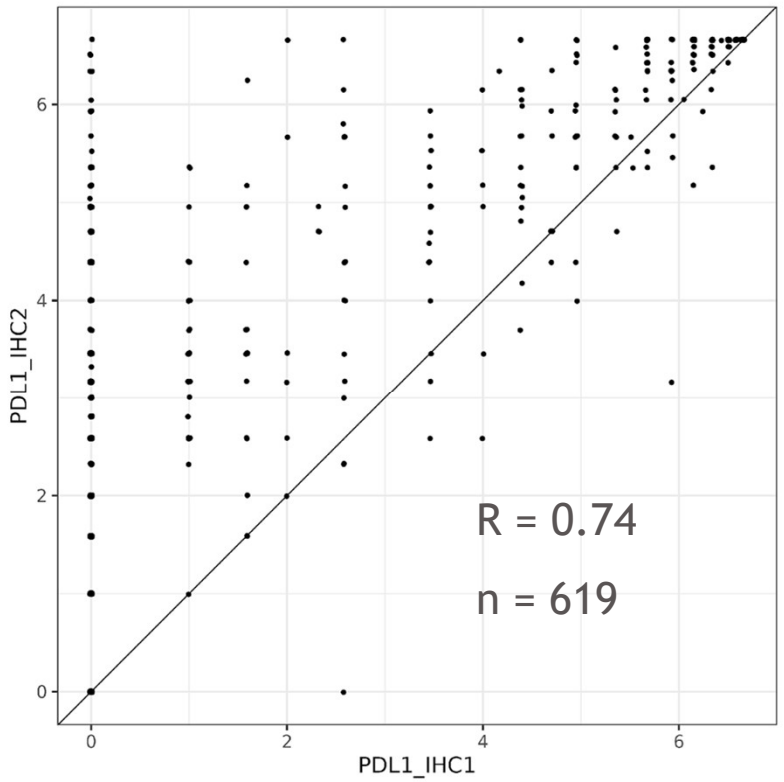
Scaling Comparison: Standard Deviation



Notes:

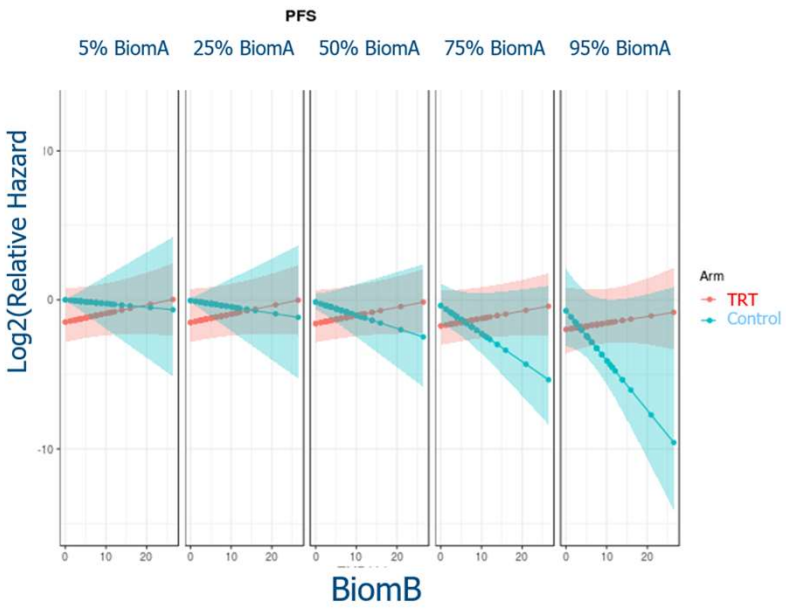
1. Used rms::rcs() to create splines because ns() has unclear scaling method.
2. We 1st scaled variables before creating basis function

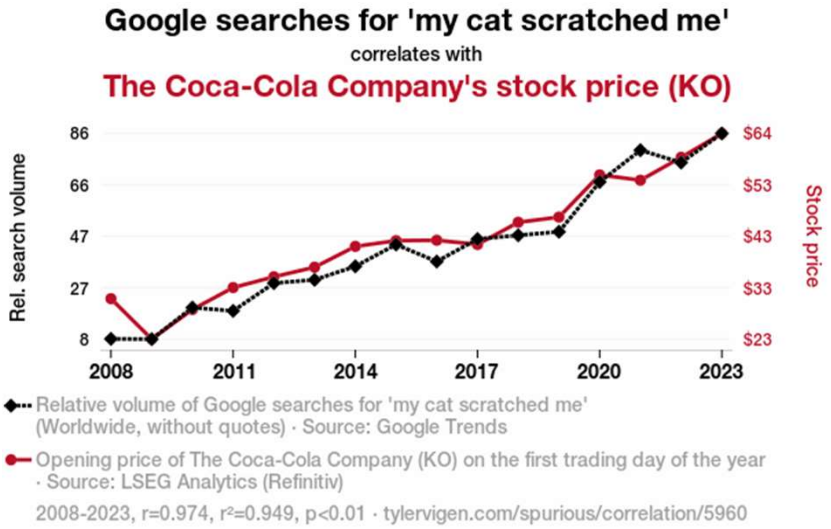
PDL1_IHC1 and PDL1_IHC2 Correlation?
(All observations)

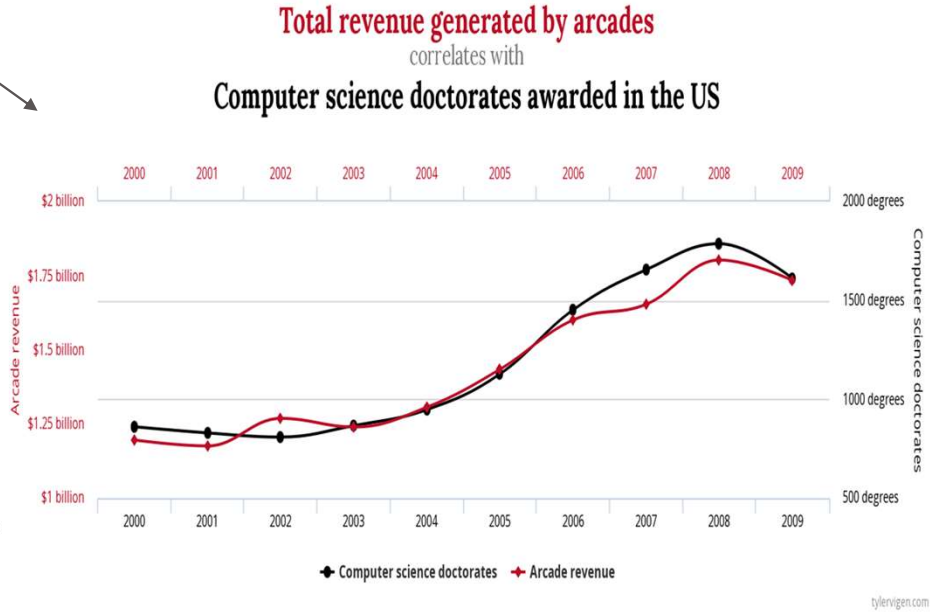


Why Multivariable Models?

1. Controls for confounders
   - Fewer spurious findings
2. Can increase precision
3. Interactions

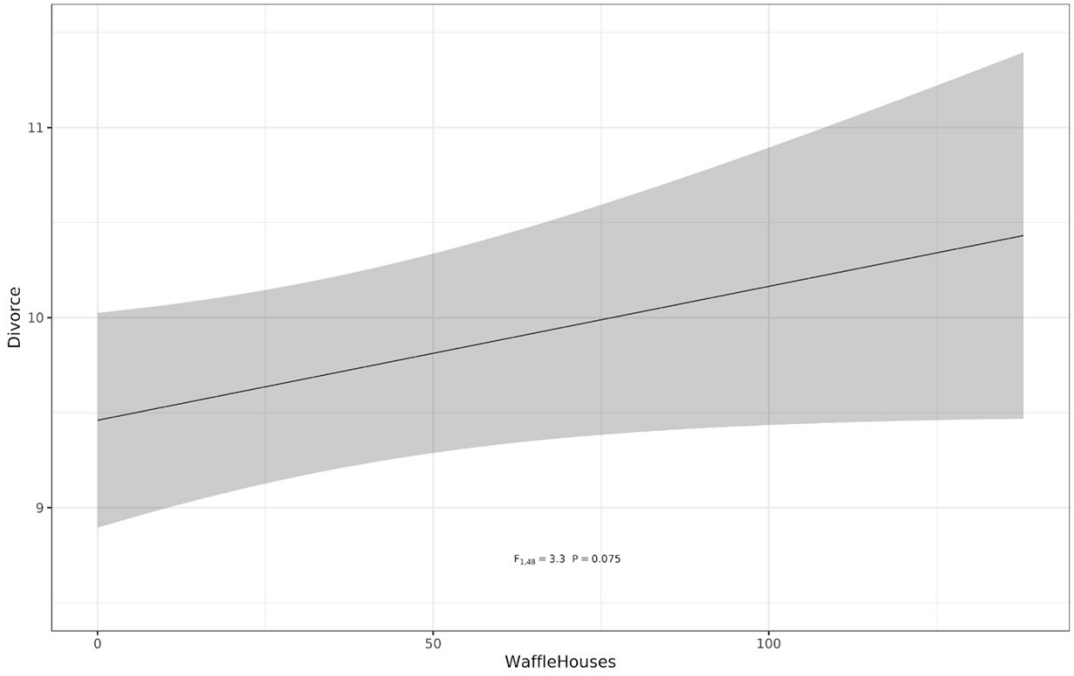






WaffleHouses ⟶ High divorce rates?

Divorce Rate ~ Waffle Houses per Million People



Well, after adjusting for location…

Divorce Rate ~ Waffle Houses per Million + South