Evaluating a facility-profiling metric based on survival probability: Application to U.S. transplant centers

Amelia Tran

April 12, 2023

Acknowledgements

- Dr. Doug Schaubel
- Dr. Peter Reese
- Dr. Ian Barnett
- Dr. Youjin Lee
- United Network for Organ Sharing (UNOS)
- National Institutes of Health (R01-DK070869)

Outline

- Background
- Motivating Example: Kidney transplantation (KT)
- Statistical Methods
  - Standardized Mortality Rate (SMR)
  - Novel Prognostic Score-Based Weighting
- Study Results
- Metric Comparison
- Discussion

Background: Facility Profiling

- Data availability on patient outcomes the last two decades
- Increased scrutiny of health care providers, especially for solid organ transplantation
- In the US, kidney transplant centers undergo two evaluations:
  - Organ Procurement and Transplantation Network (OPTN)
  - Centers for Medicare and Medicaid Services (CMS)
- Evaluations of healthcare providers and medical centers are of great interest to different parties: patients, transplant professionals and medical practitioners, etc. (Wolfe, 1994)

Motivating Example: Kidney Transplant Centers

- Post-transplant outcome by transplant center is an important factor of ensuring highest-quality care for patients
- When evaluating kidney transplant centers on survival outcomes, the most frequently used measure of mortality is standardized mortality ratio (SMR)
- Statistical methods for evaluating kidney transplant center effects:
  - Standardized mortality ratio (SMR)
  - Novel Prognostic Score-Based Weighting

Motivating Example: Kidney Transplant Centers

- Despite its wide acceptance, the SMR is not well-suited for evaluating centers due to certain limitations
  - Exaggerating center effects when survival is relatively high
  - Estimates are ill-defined if the underlying model is mis-specified
  - Indirect standardization method for averaging across case-mix covariate distributions
- Limitations of the SMR provide an inspiration to develop an alternative center effect measure: Prognostic score-based weighting method
  - Straightforward interpretation
  - Clinically self-explanatory
  - Reference population is well-defined and applied to all centers

Notation

- i: denote subject $(i = 1, 2, ..., n)$
- j: denote center $(j = 1, 2, ..., J)$
- T_i: failure time
- C_i: censoring time
- $\triangle_i = I(T_i \leq C_i)$: observed-event indicator
- $U_i = \min\{T_i, C_i\}$: observed follow-up time
- G_i: center for subject i
- $G_{ij} = I(G_i = j)$: center indicator
- X_i: observed covariate vector
- n_j: number of patients at center j
- Observed data: $O_i = (U_i, \triangle_i, X_i, G_i)$

Standardized Mortality Ratio (SMR)

- Let O_j and E_j be the observed and expected number of events at center j:

$$O_j = \sum_{i=1}^{n} G_{ij} N_i(\tau)$$

$$E_j = \sum_{i=1}^{n} G_{ij} \int_0^{\tau} Y_i(t) d\Lambda_{ij}(t)$$

- The center-specific SMR has the structure:

$$\text{SMR}_j = \frac{O_j}{E_j}$$

Standardized Mortality Ratio (SMR)

- SMR variance obtained with Poisson variance assumption

$$\mathrm{V}(\mathrm{SMR}_j) = V\left(\frac{O_j}{E_j}\right) = E_j^{-1}$$

$$\mathrm{V}(\log \mathrm{SMR}_j) = \frac{1}{E_j \mathrm{SMR}_j^2}$$

- Center effects determined based on normal distribution

$$Z_j = \frac{\log(\mathrm{SMR}_j)}{\sqrt{\mathrm{V}(\log \mathrm{SMR}_j)}} \sim N(0,1)$$

Defining Prognostic Score in Observational Studies

- Prognostic score originally established as an alternative to propensity score in observational studies (Hansen, 2008):
  - Little overlap in propensity score distribution among treatment groups
  - Researchers interested in removing systematic association between covariates and the outcome
- Defined as the association between observed covariates and potential outcome in the placebo or control group
- Can be used as a balancing score through subclassification, matching, or weighting in similar ways to the propensity score

Obtaining Prognostic Score from Cox regression

- With respect to kidney transplant center setup, there are many 'treatment' groups corresponding to transplant centers
- Under the assumption of equal covariate effects across centers, prognostic scores can be estimated using any center as the reference
- Prognostic score based on a semi-parametric center-stratified Cox model, where the baseline is unspecified and center-specific:

$$\lambda_{ij}(t; \mathbf{X}_i) = \lambda_{0j}(t)\exp(\beta^T \mathbf{X}_i)$$

- Estimated prognostic scores $\eta(\mathbf{X}_i) = \beta^T \mathbf{X}_i$ are continuous and can be used to construct in R risk classes through quantiles, deciles, etc.

Building Risk Classes from Prognostic Score

- In our study, we build $R = 5$ risk classes based on quintiles of $\eta(\mathbf{X}_i)$
- Let $Q_i = r$ denote risk class membership where $r = 1, ..., 5$
- $P(Q_i = r) = 0.2$ for all r and $Q_{ir} = I(Q_i = r)$
- Weight for each subject is then constructed:

$$\hat{w}_{ijr} = G_{ij} Q_{ir} \frac{n_j}{n_{jr}} \hat{p}_r$$

where $\hat{p}_r = n^{-1} \sum_{i=1}^{n} Q_{ir}$ and $n_{jr} = \sum_{i=1}^{n} G_{ij} Q_{ir}$

Developing Prognostic Score-based Center Effect Estimator

- Estimator of center-specific cumulative hazard $\widehat{\Lambda}_j(t)$:

$$\widehat{\Lambda}_j^w(t) = \sum_{r=1}^{R} \sum_{i=1}^{n} \int_0^t \widehat{\pi}_j(u)^{-1} \widehat{w}_{ijr} dN_{ijr}(u)$$

$$\widehat{\pi}_j(u) = \sum_{r=1}^{R} \sum_{i=1}^{n} \widehat{w}_{ijr} Y_{ijr}(u)$$

- Center-specific weighted survival function: $\widehat{S}_j^w(t) = \exp(-\widehat{\Lambda}_j^w(t))$
- Estimator for difference in survival probability:

$$\widehat{\tau}_j(t) = \widehat{S}_j^w(t) - J^{-1} \sum_{m=1}^{J} \widehat{S}_m^w(t)$$

for $j = 1, 2, ..., J$.

UNOS Data Description

- Evaluated U.S. kidney transplant centers with respect to 1-year graft survival (earliest of death, return to dialysis or repeat transplant)
- Data obtained from the United Network for Organ Sharing (UNOS)
- Study population: 58,353 patients who received a deceased-donor kidney transplant at age ≥ 18 between 1/1/16 and 12/31/20
- Excluded centers with < 25 transplants; $J = 201$ and center size ranged from 25 to 1,516 (median $= 325$)
- After truncating at 1 year post-transplant, 83% censoring
- Covariates: recipient age, sex, race, years between wait-listing (WL) and transplant, years on dialysis prior to WL, diabetes status, BMI, blood type, HCV, hypertension, malignancy, and Kidney Donor Risk Index (KDRI)

Estimating Centers Effects with SMR(1)

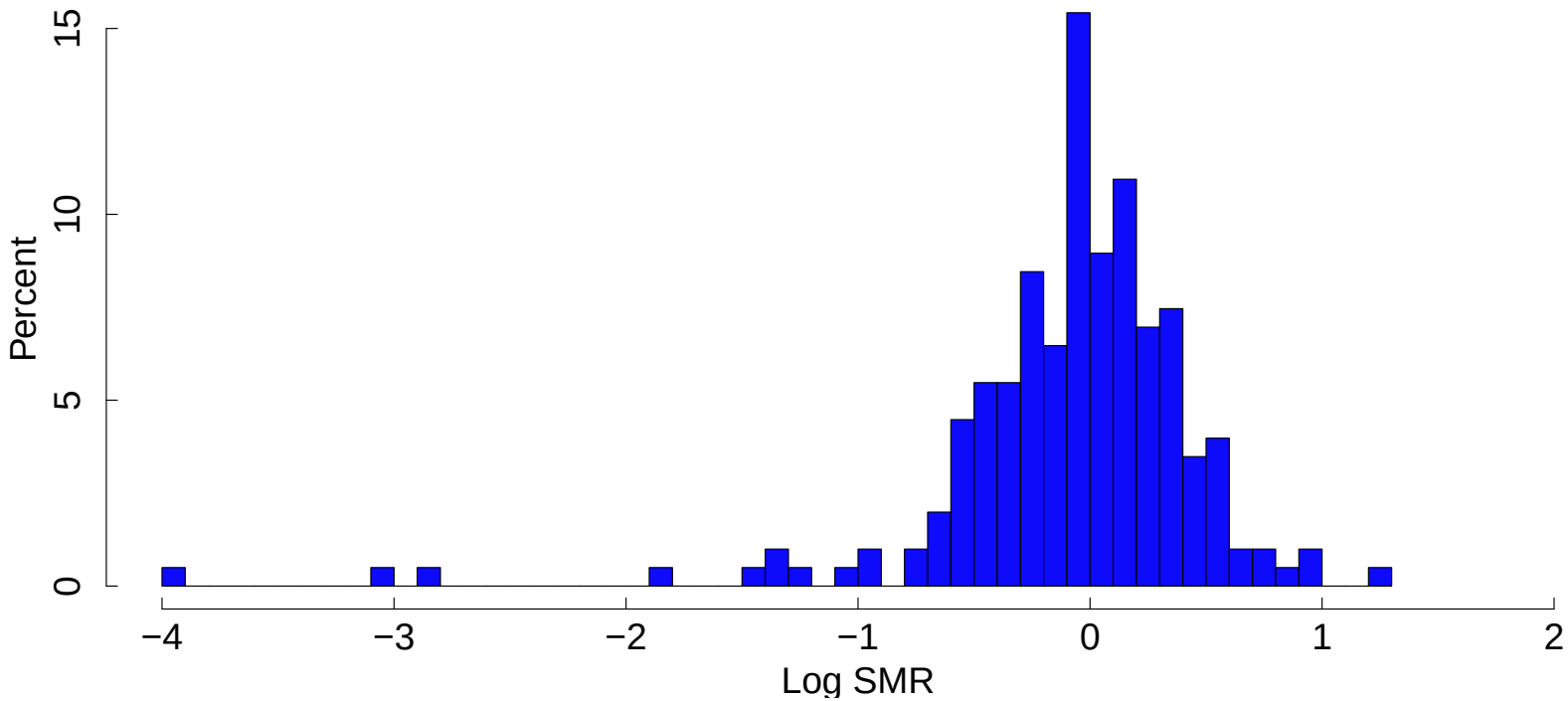

Figure 1: Histogram of log SMR in percentage

Estimating Centers Effects with $\tau(1)$

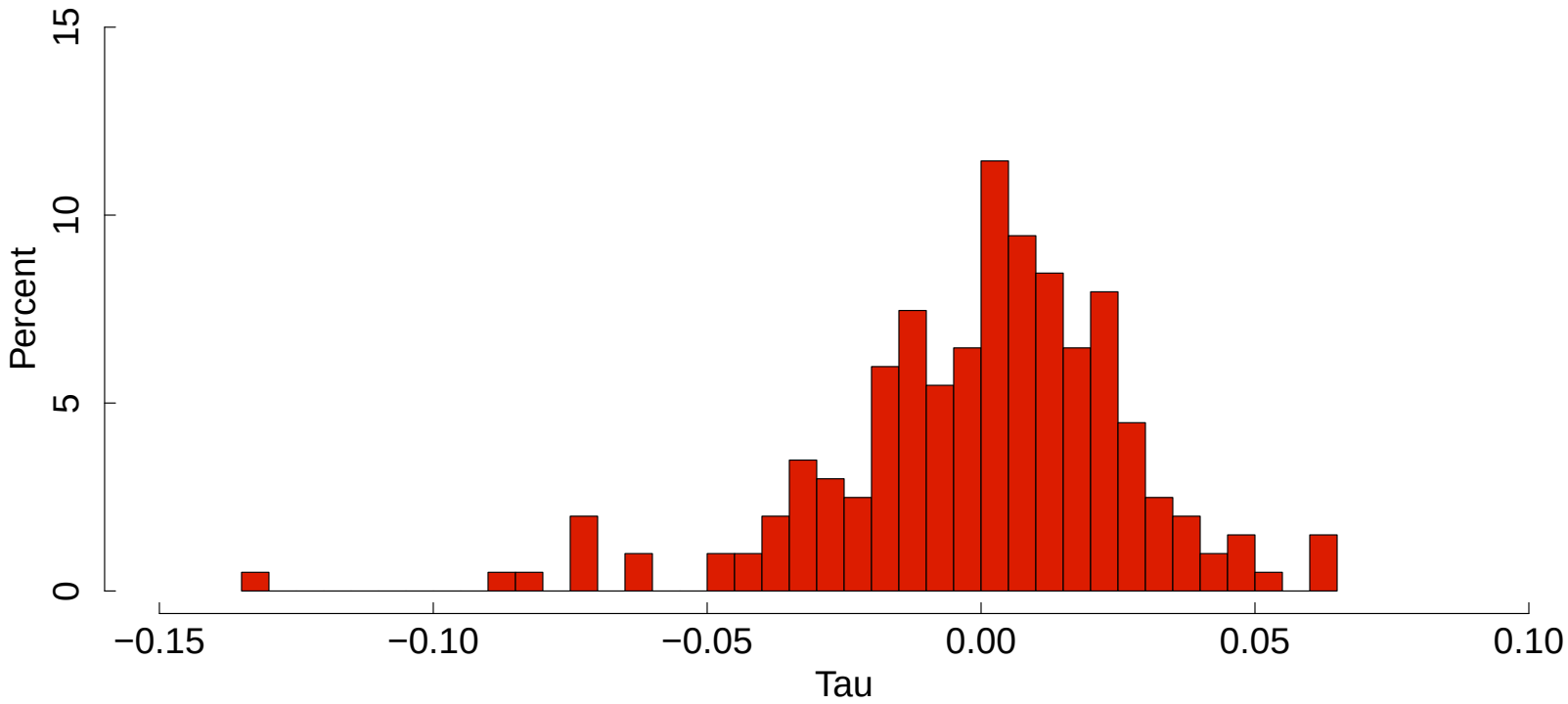

Figure 2: Histogram of excess survival τ in percentage

Estimating Centers Effects with $\tau(1)$

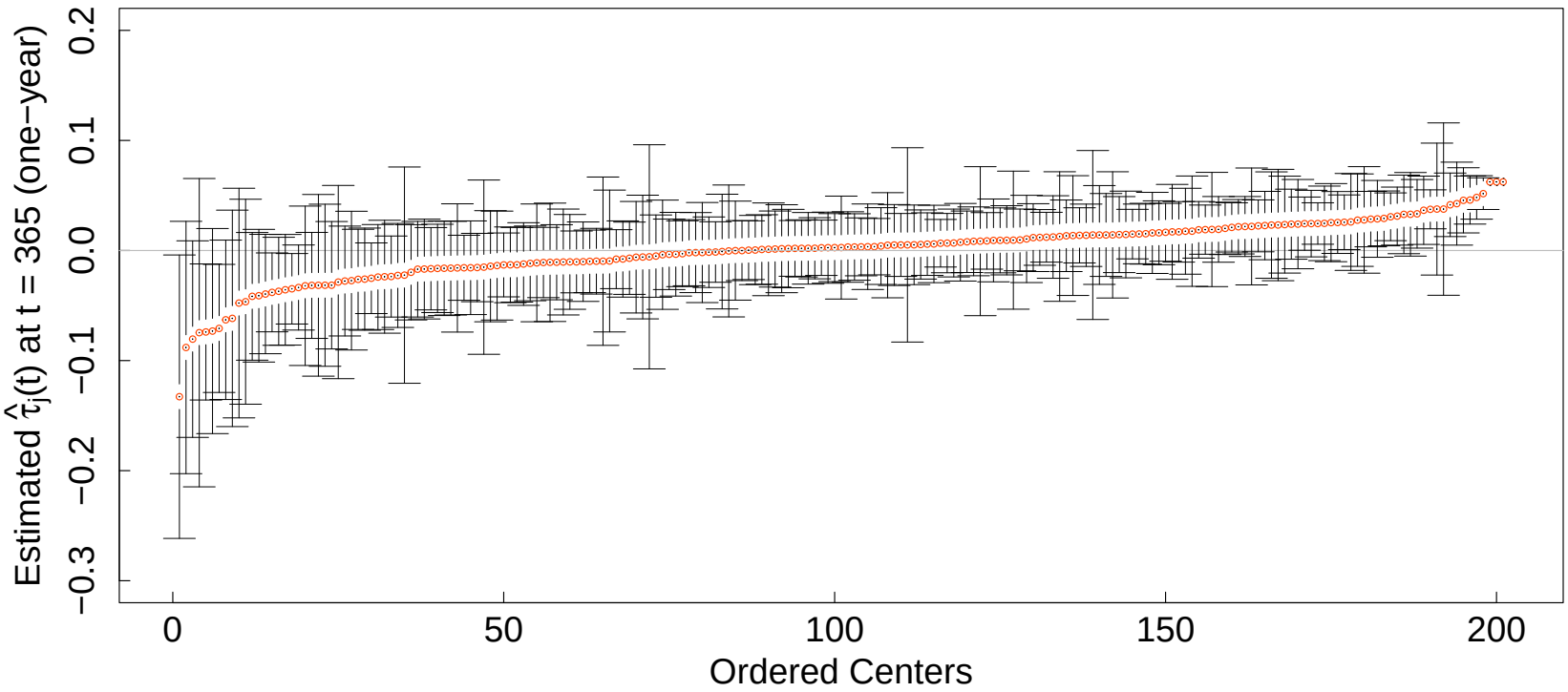

Figure 3: Excess survival probability by ordered centers

Comparison: SMR vs. Prognostic score-based weighting

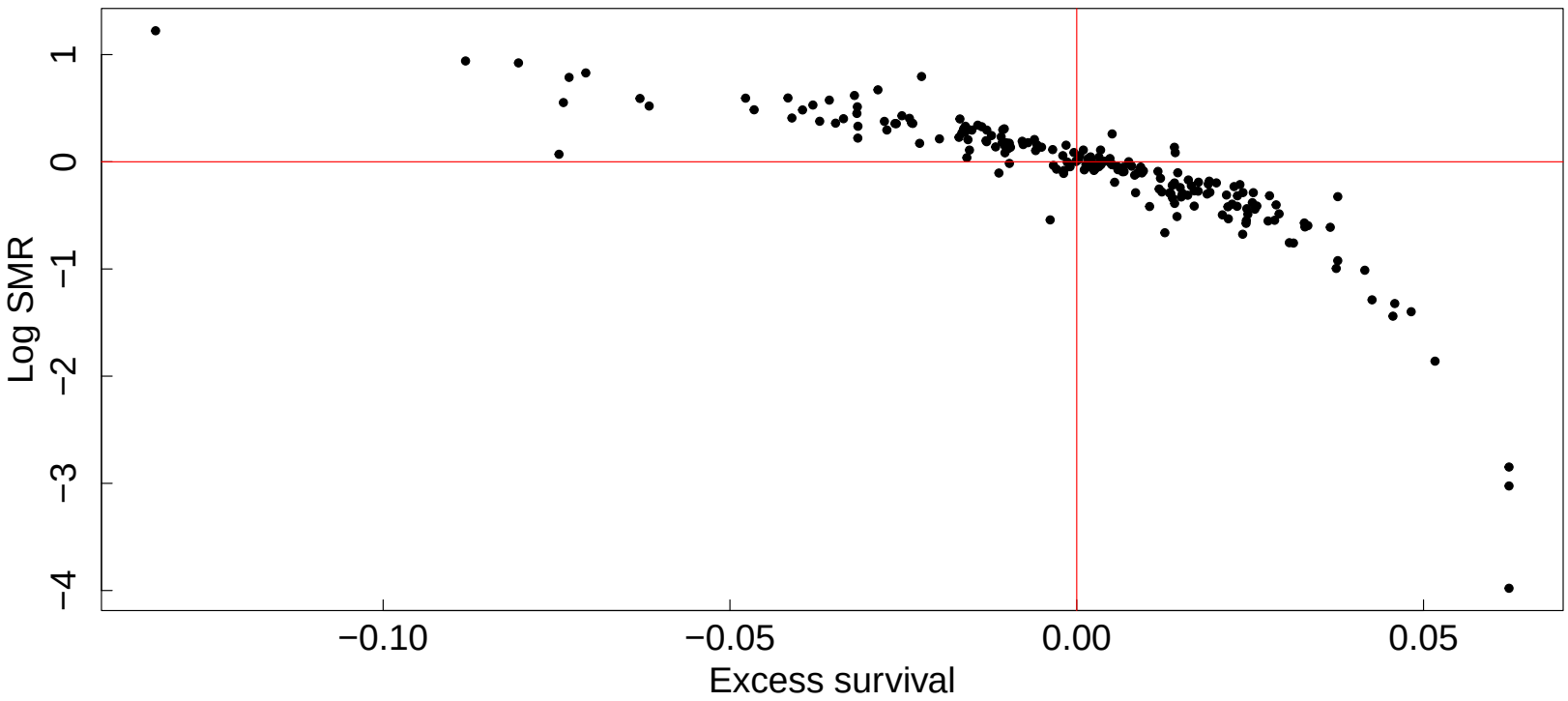

Figure 4: Scatterplot of excess survival probability and log SMR

[1]Sample correlation is -0.94

Comparison: SMR vs. Prognostic score-based weighting

Cross classification		$\tau(1)$			
		Better Center	**Null Center**	**Worse Center**	**Total**
SMR(1)	**Better**	**1**	0	0	1
	Null	18	**159**	0	177
	Worse	0	18	**5**	23
	Total	19	177	5	**201**

Table 1: Numbers of centers in each stratum

[2]Metric agreement is 82%

Comparison: SMR vs. Prognostic score-based weighting

Median SMR Median $\tau(1)$		$\tau(1)$		
		Better Center	**Null Center**	**Worse Center**
SMR(1)	**Better**	**0.37** **0.04**	NA NA	NA NA
	Null	**0.416** **0.04**	**0.97** **0.003**	NA NA
	Worse	NA NA	**1.70** **-0.03**	**1.74** **-0.07**

Table 2: Center-specific median SMR and median τ

Discussion

- We evaluated U.S. kidney transplant centers with respect to 1-year graft survival with SMR and novel prognostic score-based approach
- Correlation between the two metrics is approximately -0.94
- Novel prognostic score-based weighting method:
  - Clinically straightforward interpretation
  - Robust to model mis-specification
  - Fair facility profiling metric across all medical centers
- Would yield more accurate facility profiling in kidney transplantation
- Potential future work includes evaluating independent censoring assumption and covariate-by-center interaction

Selected References

Hansen, B. B. (2008). The prognostic analogue of the propensity score. *Biometrika*, 95(2):481–488.

Leacy, F. P. and Stuart, E. A. (2014). On the joint use of propensity and prognostic scores in estimation of the average treatment effect on the treated: a simulation study. *Statistics in medicine*, 33(20):3488–3508.

Lee, Y. and Schaubel, D. E. (2022). Facility profiling under competing risks using multivariate prognostic scores: Application to kidneytransplant centers. *Statistical Methods in Medical Research*, 31(3):563–575.

Wolfe, R. A. (1994). The standardized mortality ratio revisited: improvements, innovations, and limitations. *American Journal of Kidney Diseases*, 24(2):290–297.

Thank You!

Questions?

Email: tran26h@upenn.edu

Discussion

Summary of Logistic Regression			
Characteristics	**Odds Ratio**	**95% Confidence Interval**	**p-value**
Median # patients	1.00	(1.00, 1.00)	0.13
Median risk score	14.8	(0.47, 499)	0.13
Median KDRI	0.17	(0.00, 11.6)	0.4

Table 3: Logistic Regression on Metric Agreement

- 36 medical centers categorized differently between the two metrics
- Logistic regression to see how much impact the number of patients, median risk scores, and median KDRI have on metric agreement
- No covariates found statistically significant