Exploring Treatment Effect Heterogeneity of Right Heart Catheterization in the SUPPORT Dataset with Time-To-Event Outcomes

Abhroneel Ghosh[1], Leo Hickey[2], Quinn Hungerford[3], Isabella Summe[4]

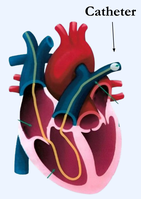
Indian Statistical Institute[1], Vassar College[2], University of Texas at Austin[3], University of Chicago[4]

Yale SCHOOL OF PUBLIC HEALTH

BIGDATA SUMMER IMMERSION

Introduction

Background



Right heart catheterization (RHC) is a procedure used to monitor cardiac function. Its effectiveness remains debated: Connors et al.'s[1] 1996 analysis of the SUPPORT dataset, looking at 5,735 critically ill patients across five U.S. hospitals, reported a negative average treatment effect. However, this may obscure subgroup-level variation in treatment effectiveness.

Research Objective

Expanding on Connors et al., we examine whether modern machine learning methods for causal inference with time-to-event data can uncover treatment effect heterogeneity (HTE) in RHC use to inform personalized care. Specifically, we ask:

> **How can the implementation of time-to-event analysis techniques explore heterogeneity in the causal effect of RHC and identify clinically significant patient subpopulations?**

Assumptions

Causal Inference Assumptions[2]:

- Positivity
- Consistency
- Exchangeability
- Independence of Observations

Survival Analysis Assumptions[2]:

- Independence of Censoring
  - 3.2% right censored (discharged and ceased contact before cutoff)
  - Tested using imputed dataset
- Proportional Hazards

Causal Estimand and Outcome

Conditional Average Treatment Effects (CATEs) are estimated using the difference in Restricted Mean Survival Time (RMST) up to a fixed time horizon (180 days).

$$\theta^{i}_{CATE} = \mathbb{E}[(T(1) \wedge \tau)|X_i] - \mathbb{E}[(T(0) \wedge \tau)|X_i], \quad \tau = 180$$

RMST reflects the expected survival time up to day τ under treatment vs. control.

We selected RMST as our outcome because it is causally interpretable, consistent across time horizons, and effective at integrating survival information over time[3]. We selected a 180-day horizon based on clinical relevance and performance.

Time Horizon	30 day	90 day	180 day
Average Correlation b/w Outcomes	0.656	0.7778	0.7915
% Observed Death	33.4%	43.7%	49.3%

Table 1: Comparison of Time Horizons

Model Overview

	DeepSurv[4]	Causal Survival Forest (CSF)[5]
Goal	Predict survival probabilities	Estimate heterogeneous treatment effect (HTE)
Method	Neural network extension of Cox Regression model	Survival tree-based extension of random forests
Focus	**Designed to have strong predictive power**	**Splits to maximize treatment effect differences**
Limitations	Assumes proportional hazards; less interpretable	Does not predict survival or have measurable performance indicator

Methods

Model Evaluation

Model	Mean C-Index
IPW-Cox with interaction terms	0.6860
DeepSurv	0.7095

Table 2: DeepSurv Model Performance

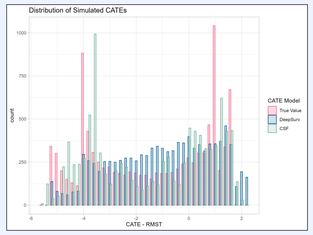
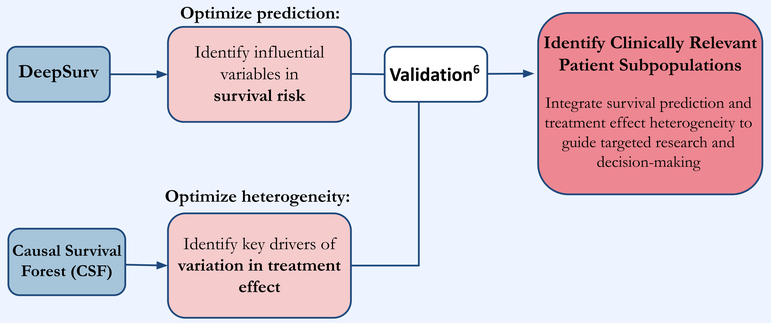


Figure 1: Comparative CATE Distributions from Simulation Study.

- **DeepSurv** has greater predictive performance than traditional methods
- Single sample of 10,000 individuals with a linear HTE simulated
- Both models perform well and align with each other, **CSF** more closely matches actual distribution

Framework for Identifying Heterogeneity



Results

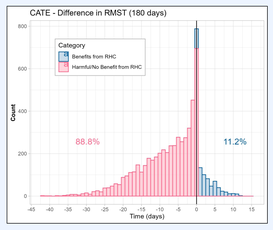


Figure 2: DeepSurv RMST 180-Day CATE Distribution

- Most patients showed little or no benefit from RHC
- Majority had a negative effect of RHC
- Small subpopulation that benefits
- Left skewed

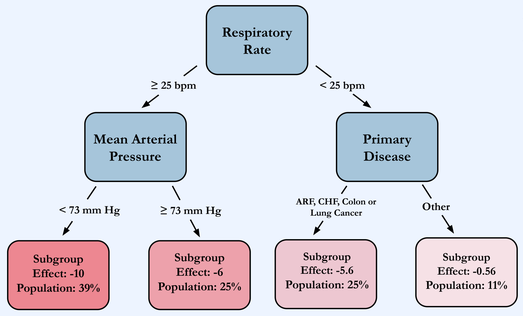


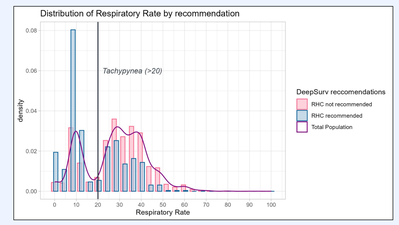
Figure 3: Decision Tree Subgroup Identification

- Decision tree model fit to regress estimated CATE values
- Simpler model for interpretability of results
- Identifies respiratory rate and mean arterial pressure as important

Conclusions

Discussing Key Covariates

1. Respiratory Rate



2. Mean Arterial Pressure

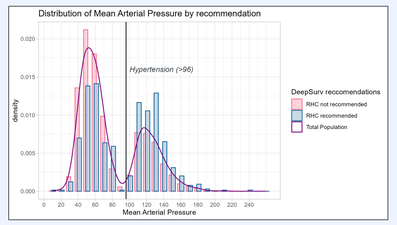


Figure 4 and 5: Distributions of Key Covariates by DeepSurv Treatment Recommendation

- Clear bimodal populations with clinically meaningful boundaries
- Populations with estimated positive/negative effect of RHC have different distributions - may be indicative of underlying medical interactions

CSF Validation and Implications

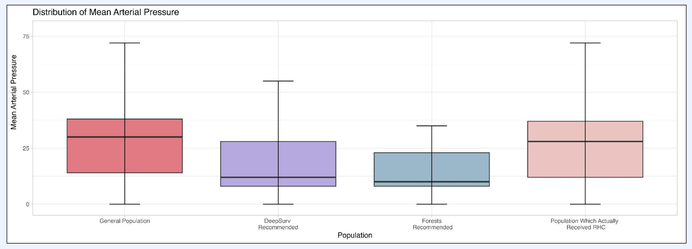


Figure 6: Mean Arterial Pressure Quantile Comparisons Across Models

Both DeepSurv and CSF identified respiratory rate and mean arterial pressure as key covariates, yielding treatment-recommended subgroups with distributions that vary notably from the overall sample and treated populations. This discrepancy highlights a potential mismatch between observed treatment assignments and model-informed recommendations. Clinicians should carefully consider these biomarkers as they individualize RHC interventions.

Limitations and Future Scope

Limitations	Future Scope
No uncertainty estimates	Exploring AFT models (AFT-BART) with non-proportional hazards. May lead to better performance and further strengthen results
Evidence of deviation from proportional hazards assumption	Validating findings on heterogeneity with external/recent data
SUPPORT exclusion criteria imply left-truncation	Further investigating the medical significance of key covariates and expanding into a validated treatment recommendation system

For their invaluable support, we would like to thank our principal investigators Dr. Fan Li and Dr. Lee Kennedy-Shaffer, program director Dr. Bhramar Mukherjee, and the BDSY staff including Xi Fang and Jiaqi Tong.

References

[1]Connors, Alfred et al. The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients. SUPPORT Investigators. JAMA, 1996.
[2]Voinot, Charlotte et al. Causal Survival Analysis, Estimation of the Average Treatment Effect (ATE): Practical Recommendations. 2025.
[3]Stensrud, Mats et al. Limitations of Hazard Ratios in Clinical Trials. European Heart Journal, 2019.
[4]Katzman, Jared et al. DeepSurv: Personalized Treatment Recommender System Using a Cox Proportional Hazards Deep Neural Network. BMC Med Res Methodol, 2018.
[5]Cui, Yifan et al. Estimating Heterogeneous Treatment Effects With Right-Censored Data via Causal Survival Forests. Journal of the Royal Statistical Society, 2023.
[6]Noroozizadeh, Shahriar et al. The Impact of Medication Non-adherence on Adverse Outcomes: Evidence from Schizophrenia Patients via Survival Analysis. 2025.