Stepwise Fine and Gray: Subject-Specific Variable Selection Shows When Hemodynamic Data Improves Prognostication of Comatose Post-Cardiac Arrest Patients

Xiaobin Shen[1], Jonathan Elmer[2], George H. Chen[1]
[1]Heinz College of Information Systems and Public Policy, Carnegie Mellon University, [2]Department of Emergency Medicine, University of Pittsburgh

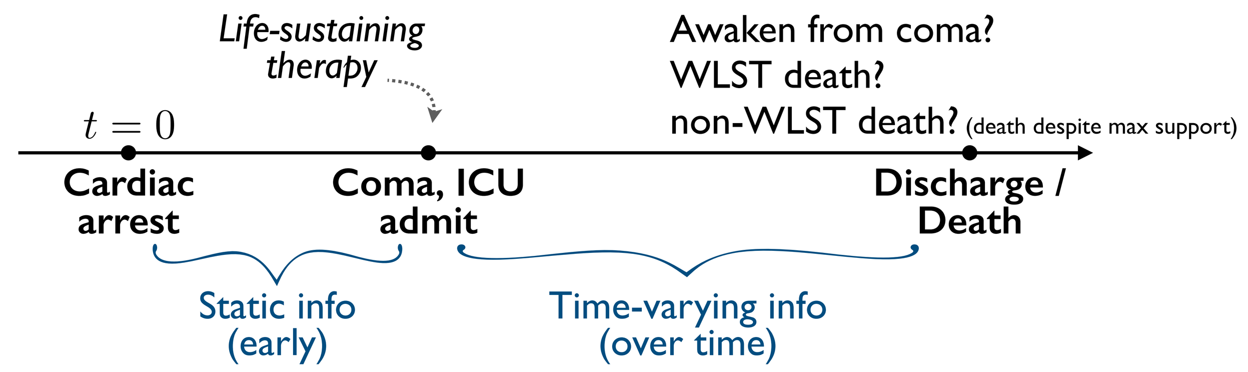
Full Paper

MACHINE LEARNING FOR HEALTH CARE
August 15 - 16, 2025

Funding: NSF CAREER award #2047981
NIH/NINDS 5R01NS124642

Background & Motivation

Comatose post-cardiac arrest patients who died from withdrawal of life-sustaining therapies (WLST) may have recovered if kept on life-sustaining therapies



Clinical information that informs **neurological prognostication** collected serially overtime

- Early **static features**: demographics, early neuro exams
- Later **dynamic features**: continuous hemodynamic data (MAP, vasopressor doses)
- Question: ***When*** and for ***whom*** does dynamic hemodynamic data add prognostic value?

Main Contribution

- **Stepwise Fine and Gray**: a competing risks model with neural nets
- Splits prediction into two phases:
  - **Phase 1**: Static features.
  - **Phase 2**: Adds time-varying hemodynamic data.
- Learns a **patient-specific threshold** to decide whether Phase 2 data improves predictions.

Competing Risks & Fine and Gray Primer

Dynamic Competing Risks Problem Setup

Cumulative Incidence Function (CIF)

$$F_{k,t}(h|x) \triangleq \mathbb{P}(T \le h, D = k \mid X_t = x)$$

x : patient feature vector
t : how much time has elapsed so far
h : time horizon (measured starting from t)
k : event (awaken, non-WLST death, WLST death)

Subhazard and CIF relationship

Define for event k the subhazard (Fine and Gray, 1999)

$\lambda_{k,t}(h|x) =$ instantaneous rate of event k happening at time horizon h given event k not yet occurring (for patient x at time t)

Proportional Subhazards Assumption

$$\lambda_{k,t}(h|x) = \lambda_{k,0}(h)\exp\{f_{k,t}(x)\} \quad \text{for all } h \ge 0, x \in \mathcal{X}$$

baseline subhazard
a (learned) risk score (e.g., neural net)

Monotonic relationship between the **CIF** and the **risk score**

increase in $f_{k,t}(x)$ → increase in $F_{k,t}(h|x)$

Interpretability of Fine and Gray

For any two feature vectors x, x', calculate the **log-subhazard ratio**

$$\ln\frac{\lambda_{k,t}(h|x')}{\lambda_{k,t}(h|x)} = f_{k,t}(x') - f_{k,t}(x)$$

Stepwise Fine and Gray

Two-Phase Decomposition

At time t after arrest, decompose feature vector: $X_t = (X_t^{(1)}, X_t^{(2)})$

static covariates elapsed time t
time-varying hemodynamics

Phase 1 uses $X_t^{(1)}$ for baseline risk

$$\lambda^{(1)}_{k,t}(h \mid X_t^{(1)}) = \lambda^{(1)}_{k,0}(h)\exp\{f^{(1)}_{k,t}(X_t^{(1)})\}$$

$\to \hat{f}^{(1)}_{k,t}(X_t)$ **Phase 1 risk score: baseline prognosis**

Phase 2 adds $X_t^{(2)}$ to refine predictions, fit a second Fine and Gray model

$$\lambda^{(2)}_{k,t}(h \mid X_t^{(1)}, X_t^{(2)}) = \lambda^{(2)}_{k,0}(h)\exp\Big\{\underbrace{\hat{f}^{(1)}_{k,t}(X_t)}_{\text{treated as fixed}} + f^{(2)}(X_t^{(1)}, X_t^{(2)})\Big\}$$

$\to \hat{f}^{(2)}_{k,t}(X_t)$ **Phase 2 risk score: isolate effect from time-varying hemodynamics**

Combined additive risk score $\hat{f}^{(1)}_{k,t}(X_t) + \hat{f}^{(2)}_{k,t}(X_t)$

Log-subhazard ratio comparing Phase 2 vs Phase 1

Incremental contribution from Phase 2 time-varying features (subject-dependent)

$$I(h|X_t) \triangleq \ln\frac{\hat{\lambda}^{(2)}_k(h \mid X_t^{(1)}, X_t^{(2)})}{\hat{\lambda}^{(1)}_k(h \mid X_t^{(1)})} = \hat{f}^{(2)}_{k,t}(X_t) + \ln\frac{\hat{\lambda}^{(2)}_{k,0}(h)}{\hat{\lambda}^{(1)}_{k,0}(h)}$$

Phase 2 risk score
Correction term on baseline hazard

Threshold rule - learned $\delta_k(h)$ on validation set

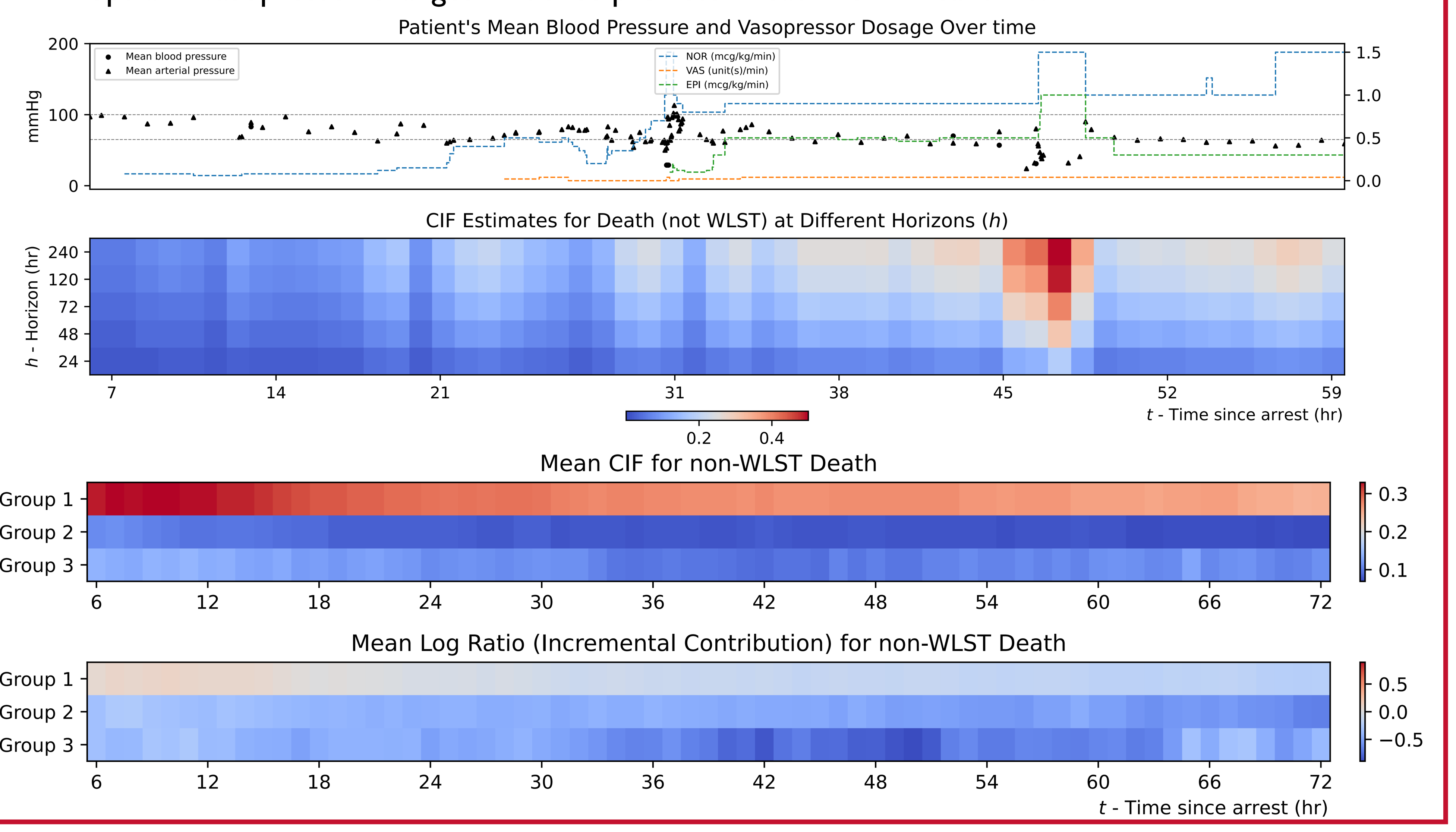
Use Phase 2 prediction only when the |incremental contribution| is larger than the threshold

if $|I(h|X_t)| \le \hat{\delta}_k(h)$, Use Phase 1 (only static features)
if $|I(h|X_t)| > \hat{\delta}_k(h)$, Use Phase 2 (time-varying features as well)

only use time-varying features when it delivers a meaningful change in predicted risk

Patient and Subgroup Visualization

MAP drops and vasopressor changes influence predicted death risk in real time.

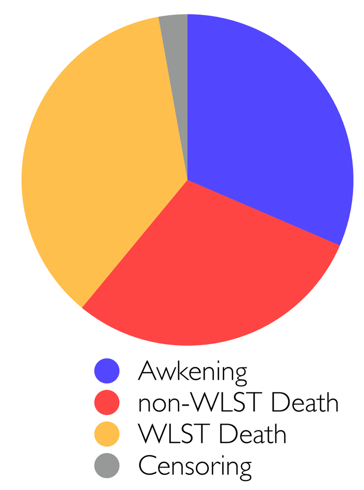


Real-World Clinical Data Experiment

Cohort

- Retrospective post-arrest ICU cohort at UPMC (2010–2022)
- 2,278 patients, with three mutually exclusive outcomes (earliest event per patient)

Proportion of the first event



Inclusion criteria

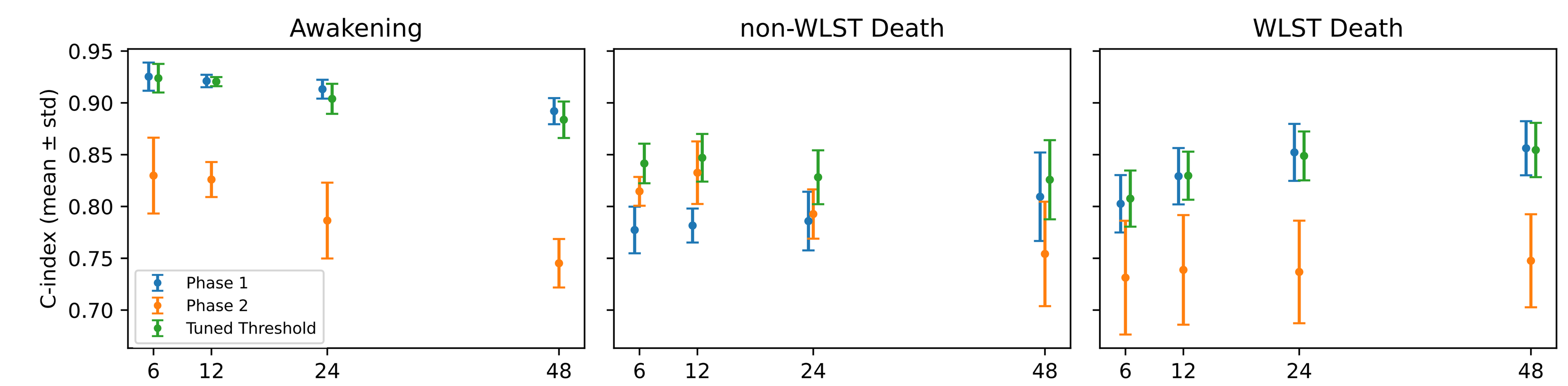
- Adult survivors of cardiac arrest, successfully resuscitated, comatose on ICU admission
- First competing event (awakening, WLST death, non-WLST death) occurred > 6 hours post-arrest (static features collected within 6hrs)

Exclusions

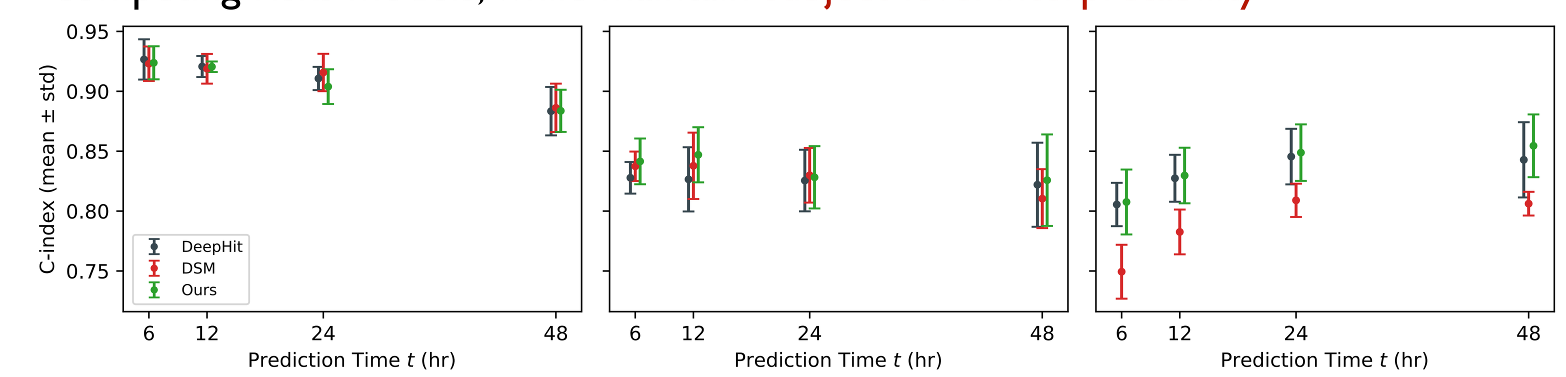
- Non-neurological WLST and events within first 6 hours (~10%)

Performance (CR c-index):

- Awakening: Static features already strong; adding hemodynamics without filtering hurts performance; thresholding recovers near-Phase-1 accuracy.
- Death (non-WLST): Hemodynamic data greatly improves prediction; thresholding further boosts performance.
- WLST: Static features dominate; hemodynamics add little.



- Competitive performance compared to recently developed neural net-based competing risks models, while maintain subject-level interpretability



Subgroup Findings - by motor component of FOUR score from early assessment

- Group 1 - Severe dysfunction: Minimal gain from hemodynamics.
- Group 2 - Moderate-mild impairment: Significant benefit from dynamic monitoring.
- Group 3 - Negative incremental contribution linked to stable/improving BP trends.

Limitation & Future Work

Conclusions & Impact

- Stepwise Fine and Gray improves prognostication by identifying when and for whom dynamic data matters
- Demonstrates the prognostic value of hemodynamic monitoring for comatose post-arrest patients
- Most value in predicting death despite maximal support, especially in patients with initial moderate-to-mild neurological impairment
- Generalizable to more phases (e.g., EEG as Phase 3) → comprehensive assessment

Limitations

- Single-center, observational data
- Limited to hemodynamic + early neuro exam features
- No advanced time-series encoders used (for interpretability)