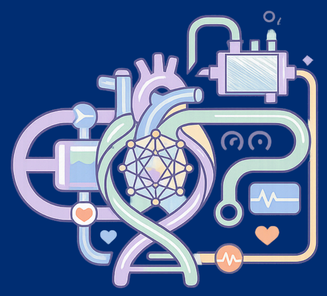


# Precision ECMO: Survival Prediction in Postcardiotomy VA-ECMO Patients


*Valeria Ventura[1], Tomisin Adebari[1], Grace Huang[1], Maria Giannakopoulos[1], Ria Thakur[1], Hanxuan Wang[1], Siyu Wang[2], Yaman Ahmad[2]*
*Joseph L. Greenstein[1], Casey Overby Taylor[1], Siamak Ardekani[1], Sung-Min Cho, DO[2]*
*[1] Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD*
*[2] Department of Neurology / Critical Care Medicine, Johns Hopkins School of Medicine, Baltimore, MD*


JOHNS HOPKINS SCHOOL of MEDICINE

JOHNS HOPKINS WHITING SCHOOL of ENGINEERING

## Introduction

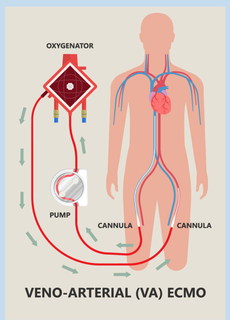
**Venoarterial extracorporeal membrane oxygenation** (VA-ECMO) is a life-saving intervention for severe cardiac and respiratory failure.

Early and accurate prognostication can guide clinical decision-making such as:

- timing of goals-of-care discussions
- escalation to durable mechanical support
- allocation of intensive care resources.

This could potentially address high in-hospital mortality rates for postcardiotomy VA-ECMO which remain 50–70% across large registries [1, 2].

Current prediction models rely on static clinical variables collected near the time of cannulation, such as the SAVE score and ENCOURAGE score [3, 4]. Growing availability of continuous bedside monitoring data presents an opportunity to leverage high-resolution hemodynamic signals for improved outcome prediction.




## Objective

**To predict survival to hospital discharge in postcardiotomy VA-ECMO patients and identify the earliest post-cannulation window at which reliable prediction is achievable.**

## Methods

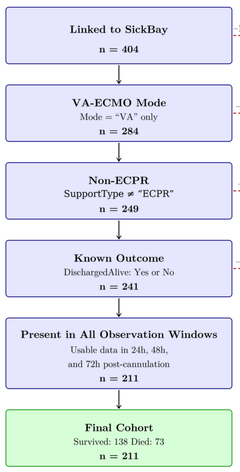



**Figure 1. Cohort selection flowchart.**
Starting from 404 SickBay-linked patients, 161 were excluded (non-VA modes, ECPR, unknown outcome), yielding 211 postcardiotomy VA-ECMO patients (138 survived, 73 died) analyzed across temporal windows.

**Step 1 — MAP Consistency Check**
If $|MAP - (DBP + \frac{SBP-DBP}{3})| \,/\, expected > 20\%$ **or** SBP ≤ DBP:
→ Set SBP, DBP, MAP to NaN for that row

**Step 2 — Raw LOCF Gap Fill**
Forward-fill gaps ≤ 15 consecutive rows (≤ 30 seconds at 0.5 Hz)
Larger gaps remain as NaN (not imputed at raw level)

**Step 3 — Derive Pulse Pressure**
$PP = SBP - DBP$, then clip PP to $[0, 200]$ mmHg
Final signal set: SBP, DBP, MAP, PP, HR, SpO2 (6 signals)

**Step 4 — Resample to 15-Minute Medians**
`series.resample('15min').median()`
24h → 96 bins  48h → 192 bins  72h → 288 bins

**Figure 2. Waveform preprocessing pipeline.** Physiologic clipping ranges applied, ABP cross-validated, Last Observation Carried Forward Fill conducted, and resampling done to ensure features were physiologically relevant and correct for the analysis.

**Study Population.** See Figure 1.

**Data Sources.** Static clinical features from the ELSO registry (Demographics, time to cannulation, lactate) were linked to continuous waveform data (SBP, DBP, MAP, HR, $SpO_2$, pulse pressure) from the SickBay clinical data warehouse.

**Preprocessing of Waveform Features.** See Figure 2.

**Feature Handling.** Correlated variables were pruned using a two-stage approach combining Spearman screening ($|\rho| > 0.85$) with physiologic domain knowledge.

**Traditional ML.** XGBoost classifiers were trained on waveform features alone (SickBay) and combined with static registry features (SickBay+ELSO) across three post-cannulation observation windows (24h, 48h, 72h), using nested 5-fold stratified cross-validation with grid-search hyperparameter tuning. A separate CatBoost classifier was trained on sequential 6-hour sub-windows for the early-prediction analysis (Figure 3).

**Deep Learning**. A 2D convolutional neural network (CNN2D) was trained on 15-minute binned waveform sequences. A hybrid variant concatenated CNN2D waveform embeddings with a dense branch on static ELSO features. Input-gradient attribution identified influential signals and time periods.

**Evaluation**. AUC-ROC with bootstrap 95% CIs; sensitivity, specificity, and F1 reported. Survival to hospital discharge was the positive class.

## Results

**XGBoost and CNN2D performance across 24h, 48h, and 72h observation windows**

| Period | Dataset | Model | AUC (95% CI) | Sensitivity | Specificity | F1 |
|---|---|---|---|---|---|---|
| First 24h | SickBay | XGBoost | 0.755 (0.686-0.821) | 0.62 | 0.80 | 0.62 |
| First 24h | SickBay+ELSO | XGBoost | 0.761 (0.691-0.825) | 0.56 | 0.77 | 0.56 |
| First 24h | SickBay | CNN2D | 0.734 (0.667-0.803) | 0.52 | 0.75 | 0.52 |
| First 24h | SickBay+ELSO | CNN2D | 0.795 (0.734-0.857) | 0.62 | 0.80 | 0.62 |
| First 48h | SickBay | XGBoost | 0.737 (0.667-0.804) | 0.58 | 0.78 | 0.58 |
| First 48h | SickBay+ELSO | XGBoost | 0.741 (0.671-0.810) | 0.59 | 0.78 | 0.59 |
| First 48h | SickBay | CNN2D | 0.761 (0.694-0.826) | 0.58 | 0.78 | 0.58 |
| First 48h | SickBay+ELSO | CNN2D | 0.797 (0.734-0.859) | 0.64 | 0.81 | 0.64 |
| First 72h | SickBay | XGBoost | 0.754 (0.682-0.822) | 0.56 | 0.77 | 0.56 |
| First 72h | SickBay+ELSO | XGBoost | 0.749 (0.678-0.816) | 0.59 | 0.78 | 0.59 |
| First 72h | SickBay | CNN2D | 0.785 (0.722-0.848) | 0.62 | 0.80 | 0.62 |
| First 72h | SickBay+ELSO | CNN2D | 0.812 (0.751-0.870) | 0.66 | 0.82 | 0.66 |

**Table 1. Best-performing model by clinical period and dataset.**

**CNN2D with combined waveform and static features** achieved the highest AUC at every observation window (0.795, 0.797, 0.812 at 24h, 48h, 72h respectively), outperforming XGBoost by a margin that widened with longer observation. XGBoost performance was stable across windows while CNN2D improved monotonically, consistent with convolutional models extracting additional temporal structure from longer sequences while summary features saturate. The single exception was the 24h SickBay-only configuration, where XGBoost (0.755) narrowly outperformed CNN2D (0.734); this reversed by 48h.

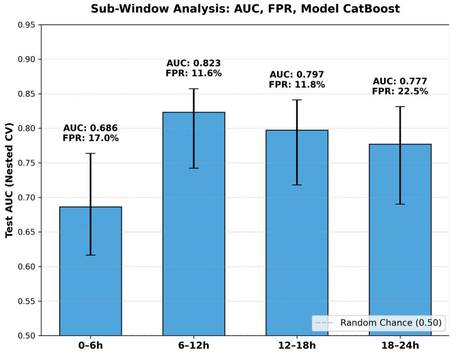



**Figure 3. Temporal optimization of ECMO mortality prediction across discrete post-cannulation sub-windows.**
Bars depict Test AUCs (with 95% CIs) across sequential 6-hour post-cannulation blocks. Annotated text displays the corresponding False Positive Rate (FPR). Following poor initial performance at 0-6h (AUC 0.686, FPR 17.0%), an **optimal threshold emerges at 6-12h**, achieving peak discrimination (AUC 0.823) and minimal false alarms (11.6%). Performance subsequently decays, with rising FPR.

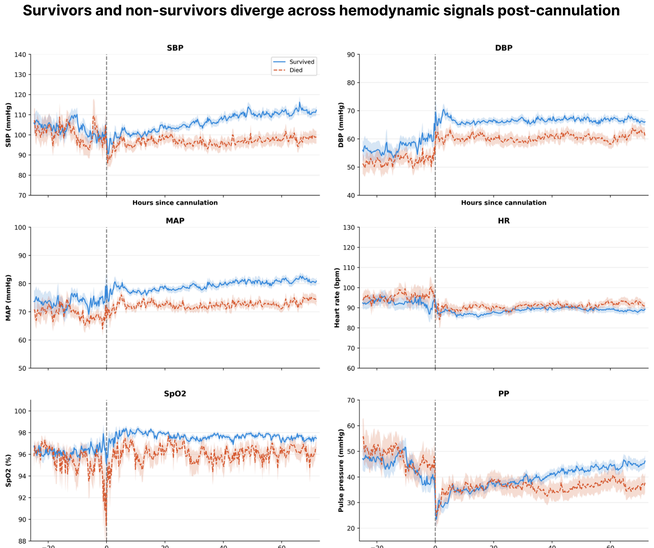



**Figure 4. Continuous hemodynamic trends around ECMO cannulation in survivors (n=138) versus non-survivors (n=73).**
Hemodynamic waveform trends after cannulation, stratified by discharge survival. Lines show the cohort mean of per-patient 15-minute medians; shaded bands show SEM. Survivors and non-survivors diverge in a staged cascade: survivors show higher mean, diastolic, and oxygen saturation within hours of cannulation; higher systolic pressure and lower heart rate emerge more gradually; and higher pulse pressure separates only after ~24h.

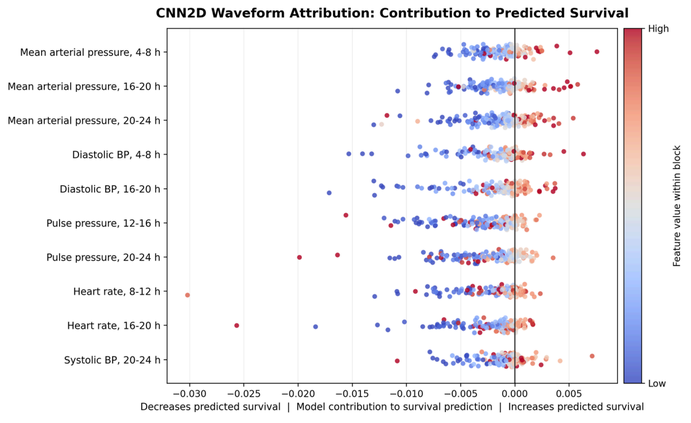



**Figure 5. Top 10 CNN2D feature contributions to 24h mortality prediction**
Top 10 CNN2D feature contributions from 24-hour post-cannulation SickBay waveforms. Each dot represents one patient. X-axis values show signed attribution toward predicted survival; negative values favor non-survival and positive values favor survival. Color indicates the feature's value relative to other patients in the same time block: red = high, blue = low.

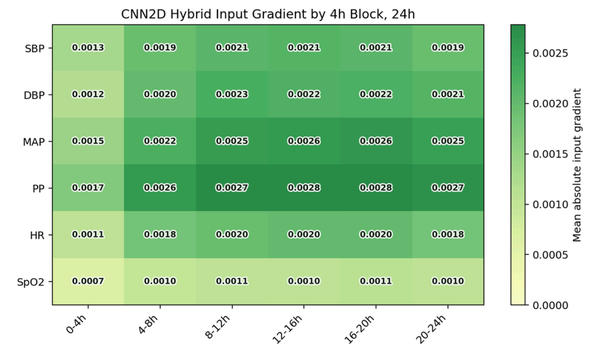



**Figure 6. Temporal feature importance for the 24-hour 2D convolutional neural network.**
Mean absolute input gradient aggregated into 4-hour blocks for each hemodynamic signal, with darker color indicating higher importance. The model showed a sustained mid-window pattern of elevated importance spanning 8–20h post-cannulation, with **peak attribution in the 12–20h period**. Pulse pressure, MAP, and DBP were the most influential signals across this window, while $SpO_2$ and HR contributed comparatively less. Importance was lowest in the immediate post-cannulation period (0–4h), suggesting the model derives minimal predictive signal from the acute cannulation transient.

## Conclusions

- **Combining continuous hemodynamic waveforms with the static ELSO registry features predicts VA-ECMO survival with good discrimination**.
- Waveform models scale with observation window while summary-feature models plateau. CNN2D AUC improved monotonically while XGBoost remained flat across the same windows, suggesting convolutional architectures extract temporal structure that 4-hour summary statistics cannot capture.
- **Early prediction is achievable** within the first 12h post-cannulation per sub-window analysis.
- Continuous waveforms **refine rather than replace static registry features**. Waveform-only CNN2D achieved moderate discrimination (AUC 0.734–0.785 across 24–72h), confirming that bedside monitoring data carries independent prognostic signal, while the hybrid model consistently outperformed either source alone.

## References


[1] Whitman GJR. Extracorporeal membrane oxygenation for the treatment of postcardiotomy shock. J Thorac Cardiovasc Surg. 2017;153(1):95-101.
[2] ELSO Registry Report. Extracorporeal Life Support Organization. 2023.
[3] Schmidt M, Burrell A, Roberts L, et al. Predicting survival after ECMO for refractory cardiogenic shock: the survival after veno-arterial-ECMO (SAVE)-score. Eur Heart J. 2015;36(33):2246-2256.
[4] Du CH, Glick D, Tung A. Error-checking intraoperative arterial line blood pressures. J Clin Monit Comput. 2019 Jun;33(3):407–412. doi: 10.1007/s10877-018-0167-7. PMID: 29869762.