Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Leading organ-failure phenotype and subsequent new-onset atrial fibrillation in prolonged ICU stays: derivation in MIMIC-IV and external validation in eICU-CRD

Hasan Burak ISLEYEN

hasanburak.isleyen@nisantasi.edu.tr

Nişantaşı University

sevil tugrul yavuz

Cam and Sakura City Hospital

Sercan Bulut

Bağcılar Eğitim ve Araştırma Hastanesi

Fatih Kizkapan

Bağcılar Eğitim ve Araştırma Hastanesi

Cevahir Alioglu

Bezmialem Foundation University Medical Faculty Hospital

erdem sunger

Istanbul Medipol University

ertugrul okuyan

Istanbul Medipol University

Article

Keywords: atrial fibrillation, critical illness, SOFA, organ dysfunction, phenotype, external validation

Posted Date: May 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9336889/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

New-onset atrial fibrillation (NOAF) is common during critical illness, but early risk stratification remains imprecise. Organ-dysfunction trajectories have been studied in critical care, yet the association between the first dominant organ-failure phenotype and subsequent NOAF has not been defined.

Methods

We performed a retrospective derivation-validation cohort study of adult ICU stays lasting at least 72 h. The derivation cohort used MIMIC-IV and the validation cohort used eICU-CRD. A five-system leading-organ phenotype was defined by the first organ-specific SOFA subscore to reach ≥ 3 within the first 72 h (cardiovascular, respiratory, renal, hepatic, coagulation, or no failure). Neurological SOFA was excluded from the primary phenotype because of limited cross-database transportability. The primary outcome was NOAF identified after 72 h of ICU stay. Multivariable logistic regression was used in the derivation cohort, and a reduced transportable model was externally validated in eICU.

Results

The derivation cohort included 11,735 ICU stays and 367 NOAF events after 72 h (3.1%). The validation cohort included 42,942 stays and 2,216 pragmatic AF events after 72 h (5.2%). In MIMIC-IV, crude event rates were highest in the cardiovascular-leading phenotype (5.5%) and respiratory-leading phenotype (3.3%), compared with no failure (1.4%). In the primary adjusted derivation model, cardiovascular-leading phenotype was associated with higher odds of NOAF versus no failure (OR 1.81, 95% CI 1.01–3.23; p = 0.047), whereas the respiratory-leading phenotype was not independently significant (OR 1.32, 95% CI 0.86–2.03; p = 0.206). In eICU, crude event rates were again highest in cardiovascular-leading (7.5%) and respiratory-leading phenotypes (5.7%), compared with no failure (4.2%). When the reduced transportable derivation model was applied to eICU, discrimination was modest (AUROC 0.68), calibration slope was 0.64, and calibration intercept was 1.75. In the eICU refit model, cardiovascular-leading phenotype remained associated with higher odds of pragmatic AF after 72 h (OR 1.20, 95% CI 1.04–1.40; p = 0.015).

Conclusions

A clinically transparent five-system leading-organ phenotype identified differential NOAF risk across prolonged ICU stays. Cardiovascular-leading phenotype showed the most consistent signal across derivation and validation cohorts. This approach may complement existing critical-care NOAF risk models by adding a temporally ordered organ-dysfunction framework.

Introduction

Atrial fibrillation is the most common sustained arrhythmia encountered in critical care. New-onset atrial fibrillation (NOAF) during critical illness is associated with hemodynamic instability, thromboembolic risk, prolonged ICU stay, and excess mortality [5–7, 21–23]. Existing reviews and consensus work suggest that NOAF in the ICU is driven by

a combination of substrate and trigger: adrenergic stress, inflammation, vasopressor exposure, volume shifts, respiratory failure, and electrolyte disturbance [5, 7–10, 20].

Most available ICU NOAF studies have focused on overall incidence, outcomes, or early multivariable risk prediction [6, 10, 11, 14–16]. In parallel, critical-care phenotype research has increasingly used SOFA trajectories to describe heterogeneous patterns of organ dysfunction [1, 13, 21]. However, these two literatures have rarely been linked. Whether the organ system that first becomes severely dysfunctional provides clinically meaningful information about subsequent NOAF risk remains unclear.

A rule-based, temporally ordered leading-organ phenotype derived from the first 72 hours of ICU stay evaluated differential risk of subsequent NOAF. A five-system phenotype captured cardiovascular, respiratory, renal, hepatic, and coagulation SOFA components. Neurological SOFA was excluded from the primary exposure because sedation- and intubation-related Glasgow Coma Scale distortion limits stable harmonization across datasets [20]. Phenotype derivation utilized MIMIC-IV, with subsequent external validation performed in the multicenter eICU-CRD [2, 3]

Methods

Study design and reporting

We performed a retrospective derivation-validation cohort study using two deidentified public critical-care databases. MIMIC-IV served as the derivation cohort and eICU-CRD served as the external validation cohort [2, 3]. This study was reported in accordance with the STROBE statement [4].

Data sources

MIMIC-IV is a freely accessible critical-care electronic health record database derived from Beth Israel Deaconess Medical Center [2]. eICU-CRD is a multicenter ICU database containing admissions from hospitals across the United States [3]. Both resources are deidentified and require credentialed access.

Study population

We included adult ICU stays that lasted at least 72 h. This threshold was prespecified because phenotype definition required organ-specific SOFA assessment across the first 3 ICU days. We excluded stays with prior AF, postcardiotomy admissions, and pacemaker exposure during cohort construction. The derivation cohort contained 11,735 stays. The external validation cohort contained 42,942 stays.

Primary exposure

The primary exposure was a five-system leading-organ phenotype. For each ICU day during the first 72 h, organ-specific SOFA subscores were derived for cardiovascular, respiratory, renal, hepatic, and coagulation dysfunction. The leading-organ phenotype was defined as the first organ system to reach a SOFA subscore of at least 3 within the first 72 h. If more than one organ system reached the threshold on the same day, the organ with the higher subscore was selected. If scores tied, a deterministic prespecified tie-break rule was applied. Stays in which no organ system reached a subscore of at least 3 were classified as no failure. A six-system phenotype that additionally included neurological SOFA was retained only for sensitivity analysis.

Outcome definition

The primary outcome was NOAF identified after 72 h of ICU stay. In the derivation cohort, the outcome was defined using time-aware rhythm documentation after the exposure window. In the validation cohort, AF ascertainment was more pragmatic and was treated as a pragmatic AF-after-72-h endpoint. Secondary outcomes were AF before discharge, 28-day death, and in-hospital death.

Covariates

Prespecified derivation covariates were age, sex, Charlson comorbidity index, sepsis flag, day 1 SOFA total, day 3 SOFA total, potassium minimum in the first 48 h, magnesium minimum in the first 48 h, platelet minimum in the first 72 h, creatinine maximum in the first 72 h, any anticoagulation exposure in the first 72 h, heparin exposure in the first 72 h, vasopressor exposure in the first 72 h, mechanical ventilation in the first 72 h, and CRRT in the first 72 h. Potassium and magnesium were prespecified because electrolyte disturbances and magnesium-related AF literature support mechanistic relevance in atrial fibrillation [18, 19]. For cross-database transportability, the external validation model excluded variables that were structurally nontransportable or had excessive validation-cohort missingness, namely Charlson index, sepsis flag, magnesium, and heparin exposure.

Statistical analysis

Continuous variables are presented as median and interquartile range. Categorical variables are presented as number and percentage. Absolute standardized mean differences were used to describe between-group imbalance in descriptive tables.

The primary endpoint was new-onset atrial fibrillation after 72 h in the ICU. The primary analyses used multivariable logistic regression because the main endpoint was binary. In the derivation cohort, we first fit a full model that included the five-system leading-organ phenotype and all prespecified derivation covariates. We then fit a reduced transportable model using only variables that were available with stable harmonization across both databases. Odds ratios and 95% confidence intervals were reported. A secondary time-to-event analysis used Cox proportional hazards regression for time to atrial fibrillation.

External validation was performed by applying the reduced derivation-model coefficients to the eICU cohort. Model discrimination was assessed with the area under the receiver operating characteristic curve (AUROC), which represents the C-statistic for this logistic model setting. Calibration was assessed with the calibration slope and calibration intercept, and was displayed graphically with a calibration plot.

The full derivation covariates were age, sex, Charlson comorbidity index, sepsis flag, day 1 SOFA total, day 3 SOFA total, potassium minimum in the first 48 h, magnesium minimum in the first 48 h, platelet minimum in the first 72 h, creatinine maximum in the first 72 h, any anticoagulation exposure in the first 72 h, heparin exposure in the first 72 h, vasopressor exposure in the first 72 h, mechanical ventilation in the first 72 h, and CRRT in the first 72 h. For cross-database transportability, the external validation model excluded variables that were structurally non-transportable or had excessive missingness in the validation cohort.

Missingness in the covariates that entered the full derivation model was low: potassium minimum 0.50%, magnesium minimum 0.66%, platelet minimum 0.47%, and creatinine maximum 0.46%, with all other modeled variables complete. Complete-case analysis therefore yielded a final derivation cohort of 11,656 stays for the primary model and 11,675 stays for the reduced transportable model. In eICU, missingness in the reduced-model

covariates was 1.93% for potassium minimum, 2.84% for platelet minimum, and 1.36% for creatinine maximum, with all other reduced-model variables complete, yielding a validation cohort of 41,271 stays.

Continuous covariates, including potassium and creatinine, were entered into the regression models as linear terms. We did not perform a formal linearity assessment, such as spline-based modeling or a Box-Tidwell-type procedure, before fitting the final models. Accordingly, the reported models assume an approximately linear relationship on the log-odds scale for continuous predictors. This should be interpreted as a modeling limitation.

We did not formally test the missing-data mechanism and therefore did not classify missingness as MCAR or MAR. We did not apply formal multiplicity correction. Sensitivity analyses were interpreted as supportive. All analyses and figure generation were performed in Python using pandas for data management, NumPy for numerical operations, statsmodels for regression analyses, and matplotlib for figure generation.

Results

The derivation cohort included 11,735 MIMIC-IV ICU stays and the validation cohort included 42,942 eICU stays. Subsequent AF after 72 h occurred in 367 derivation stays (3.1%) and 2,216 validation stays (5.2%). In-hospital death occurred in 1,844 derivation stays (15.7%) and 5,643 validation stays (13.1%). The study flow is shown in Fig. 1.

In MIMIC-IV, the five-system phenotype distribution was no failure in 4,843 stays (41.3%), cardiovascular-leading in 3,325 (28.3%), respiratory-leading in 2,702 (23.0%), renal-leading in 409 (3.5%), coagulation-leading in 235 (2.0%), and hepatic-leading in 221 (1.9%). In eICU, the corresponding distribution was no failure in 20,662 stays (48.1%), respiratory-leading in 12,089 (28.2%), cardiovascular-leading in 5,582 (13.0%), renal-leading in 3,251 (7.6%), coagulation-leading in 951 (2.2%), and hepatic-leading in 407 (0.9%). Baseline characteristics by phenotype are presented in Table 1. The daily dominant-state pattern in the MIMIC-IV derivation cohort is shown in Fig. 2.

Table 1
Baseline characteristics in the MIMIC-IV derivation cohort by five-system leading-organ phenotype.

Variable	No failure	Cardiovascular	Respiratory	Renal	Hepatic	Coagulation	Max absolute SMD vs No failure
Age, years	63.0 [50.0, 74.0]	63.0 [51.0, 73.0]	58.5 [46.0, 69.0]	62.0 [47.0, 72.0]	54.0 [44.0, 61.0]	59.0 [48.0, 67.0]	0.525
Male sex	2470 (51.0%)	1792 (53.9%)	1558 (57.7%)	257 (62.8%)	137 (62.0%)	135 (57.4%)	0.241
Charlson comorbidity index	4.0 [2.0, 6.0]	4.0 [2.0, 6.0]	4.0 [2.0, 6.0]	6.0 [4.0, 8.0]	4.0 [3.0, 6.0]	5.0 [3.0, 7.0]	0.595
Sepsis flag	1367 (28.2%)	2376 (71.5%)	1541 (57.0%)	226 (55.3%)	137 (62.0%)	135 (57.4%)	0.959
Day 1 SOFA total	5.0 [5.0, 7.0]	12.0 [10.0, 14.0]	9.0 [8.0, 11.0]	10.0 [9.0, 12.0]	11.0 [9.0, 14.0]	10.0 [8.5, 12.0]	2.514
Day 3 SOFA total	5.0 [4.0, 6.0]	10.0 [7.0, 12.0]	8.0 [7.0, 10.0]	10.0 [8.0, 12.0]	12.0 [9.0, 14.0]	10.0 [8.0, 12.0]	2.358
Potassium minimum, first 48 h (mmol/L)	3.6 [3.4, 3.9]	3.5 [3.2, 3.9]	3.6 [3.3, 3.9]	3.8 [3.4, 4.2]	3.5 [3.1, 3.9]	3.5 [3.2, 3.8]	0.356
Magnesium minimum, first 48 h (mg/dL)	1.8 [1.7, 2.0]	1.7 [1.5, 1.9]	1.8 [1.6, 2.0]	1.9 [1.7, 2.1]	1.9 [1.6, 2.1]	1.7 [1.5, 1.9]	0.422
Platelet minimum, first 72 h (×10^9/L)	189.0 [144.0, 242.0]	128.0 [79.0, 190.0]	162.0 [110.0, 220.0]	140.0 [86.0, 197.0]	68.0 [44.0, 125.0]	27.0 [15.0, 37.0]	2.848
Creatinine maximum, first 72 h (mg/dL)	0.9 [0.7, 1.2]	1.3 [0.9, 2.4]	1.1 [0.8, 1.6]	6.4 [4.9, 8.1]	1.6 [0.9, 2.6]	1.2 [0.9, 1.9]	2.380
Any anticoagulation exposure, first 72 h	4206 (86.8%)	2939 (88.4%)	2392 (88.5%)	371 (90.7%)	185 (83.7%)	165 (70.2%)	0.414
Heparin exposure, first 72 h	4104 (84.7%)	2873 (86.4%)	2324 (86.0%)	368 (90.0%)	184 (83.3%)	162 (68.9%)	0.381
Vasopressor exposure, first 72 h	30 (0.6%)	3325 (100.0%)	448 (16.6%)	80 (19.6%)	46 (20.8%)	31 (13.2%)	17.913
Mechanical ventilation, first 72 h	897 (18.5%)	2710 (81.5%)	2702 (100.0%)	156 (38.1%)	87 (39.4%)	81 (34.5%)	2.966

Variable	No failure	Cardiovascular	Respiratory	Renal	Hepatic	Coagulation	Max absolute SMD vs No failure
CRRT, first 72 h	13 (0.3%)	273 (8.2%)	80 (3.0%)	83 (20.3%)	20 (9.0%)	8 (3.4%)	0.698

Crude derivation event rates were lowest in the no-failure group (1.4%) and highest in the cardiovascular-leading group (5.5%). Respiratory-leading (3.3%), coagulation-leading (3.4%), hepatic-leading (3.2%), and renal-leading (2.9%) phenotypes showed intermediate event rates. Validation-cohort crude rates showed the same general ranking pattern, with the highest event rate in the cardiovascular-leading phenotype (7.5%), followed by coagulation-leading (6.1%), respiratory-leading (5.7%), renal-leading (5.3%), no failure (4.2%), and hepatic-leading (2.7%). Crude phenotype-specific event rates are summarized in Table 3 and illustrated in Fig. 3.

Table 3
Crude incidence of subsequent atrial fibrillation by phenotype in derivation and validation cohorts.

Phenotype	MIMIC N	MIMIC events	MIMIC rate %	eICU N	eICU events	eICU rate %
No failure	4843	70	1.4	20662	865	4.2
Cardiovascular	3325	182	5.5	5582	418	7.5
Respiratory	2702	88	3.3	12089	693	5.7
Renal	409	12	2.9	3251	171	5.3
Hepatic	221	7	3.2	407	11	2.7
Coagulation	235	8	3.4	951	58	6.1

In the full derivation model, cardiovascular-leading phenotype was independently associated with higher odds of subsequent NOAF compared with no failure (OR 1.81, 95% CI 1.01–3.23; p = 0.047). The respiratory-leading phenotype was directionally associated with higher odds but did not reach statistical significance (OR 1.32, 95% CI 0.86–2.03; p = 0.206). Renal-leading, hepatic-leading, and coagulation-leading phenotypes were not independently associated with the primary endpoint. The reduced transportable derivation model produced similar estimates, with cardiovascular-leading phenotype remaining associated with higher odds of NOAF (OR 1.81, 95% CI 1.01–3.25; p = 0.046). Adjusted phenotype associations in derivation and validation cohorts are summarized in Table 2 and Fig. 4.

Table 2
Association between five-system leading-organ phenotype and subsequent atrial fibrillation.

Comparison	MIMIC adjusted OR	MIMIC adjusted CI low	MIMIC adjusted CI high	MIMIC adjusted P	eICU adjusted OR	eICU adjusted CI low	eICU adjusted CI high	eICU adjusted P
Cardiovascular vs No failure	1.81	1.01	3.23	0.047	1.20	1.04	1.40	0.015
Respiratory vs No failure	1.32	0.86	2.03	0.206	1.08	0.94	1.25	0.263
Renal vs No failure	0.62	0.30	1.28	0.199	0.89	0.70	1.12	0.323
Hepatic vs No failure	0.73	0.31	1.74	0.474	0.54	0.29	1.00	0.050
Coagulation vs No failure	1.00	0.44	2.24	0.991	1.19	0.88	1.60	0.254

When the reduced derivation model was transported to eICU, discrimination was modest (AUROC 0.68) and calibration showed underprediction in the validation cohort (slope 0.64; intercept 1.75; Brier score 0.051). In the eICU refit reduced model, cardiovascular-leading phenotype remained associated with higher odds of pragmatic AF after 72 h (OR 1.20, 95% CI 1.04–1.40; p = 0.015). Respiratory-leading phenotype was not independently significant after adjustment (OR 1.08, 95% CI 0.94–1.25; p = 0.263). Model performance metrics are summarized in Table 4, and calibration of the transported reduced model is shown in Fig. 5.

Table 4
Model performance in derivation and external validation.

Model	N	AUROC	Brier score	Calibration slope	Calibration intercept
Derivation full model (MIMIC)	11656	0.803	0.029		
Derivation reduced transportable model (MIMIC)	11675	0.797	0.029		
External validation of transported reduced model (eICU)	41271	0.683	0.051	0.642	1.755
eICU refit reduced model	41271	0.688	0.049		

Secondary outcome patterns were directionally coherent. In MIMIC-IV, hospital mortality was 7.4% in the no-failure group, 23.3% in cardiovascular-leading stays, 16.9% in respiratory-leading stays, 21.5% in renal-leading stays, 40.3% in hepatic-leading stays, and 32.3% in coagulation-leading stays. In eICU, hospital mortality was 7.6%, 19.4%, 18.0%, 14.0%, 31.7%, and 25.1% across the same phenotype categories.

Discussion

A five-system leading-organ phenotype derived from the first 72 hours of ICU admission identified differential risk of subsequent atrial fibrillation across two independent databases. The cardiovascular-leading phenotype demonstrated the most consistent association with later AF, maintaining significance after multivariable adjustment. Temporally ordered organ-failure patterning captures early arrhythmogenic substrate prior to overt

arrhythmia. Bedford et al. previously outlined triggers for NOAF during critical illness, highlighting the transition from early physiologic stress to sustained arrhythmias [8, 9].

Cardiovascular-leading stays exhibited the highest crude AF rates in the derivation (5.5%) and validation (7.5%) cohorts, retaining an independent association with NOAF after severity adjustment. Early dominant cardiovascular dysfunction summarizes an arrhythmogenic state extending beyond age, standard SOFA burden, or baseline renal dysfunction. Acute cardiovascular stress drives adrenergic surges, hemodynamic instability, and atrial stretch. Bosch et al. identified vasopressor exposure and systemic inflammation as primary triggers for NOAF in a meta-analysis of critical care patients [10].

Respiratory-leading phenotype stays demonstrated elevated crude NOAF rates (3.3% in derivation, 5.7% in validation) compared to stays with no failure, though the independent association attenuated after multivariable adjustment. Hypoxemia, positive-pressure ventilation, and acid-base disturbances generate a well-recognized trigger milieu for atrial arrhythmias. Respiratory dysfunction drives arrhythmia risk primarily through correlated illness severity and treatment intensity rather than an isolated electrophysiologic effect. Walkey and colleagues previously linked respiratory failure to NOAF, demonstrating that mechanical ventilation significantly amplifies arrhythmic risk [5].

Rule-based phenotyping anchored to the first organ system crossing a predefined dysfunction threshold preserved temporal ordering and supported cross-database harmonization. Trajectory-clustering models typically evaluate total SOFA progression, often obscuring the specific index organ failure driving subsequent clinical events. Identifying the initial dominant organ failure isolates the earliest physiologic divergence. Xu et al. applied trajectory-based subphenotyping to sepsis, demonstrating that distinct temporal patterns of organ dysfunction carry diverging prognostic implications [13].

The transported reduced model maintained above-chance discrimination (AUROC 0.68) in the eICU cohort, while the calibration slope (0.64) and positive intercept (1.75) indicated underprediction of overall event probability [17]. Prediction transport across critical care databases frequently encounters calibration decay when phenotype definitions require harmonized organ-support variables. The refitted eICU model successfully reproduced the primary cardiovascular-leading signal. Stevens and Poppe demonstrated that calibration slope shifts reflect differences in predictor distributions and baseline risks across disparate clinical populations [17].

Inclusion of the neurological SOFA component in preliminary derivation models dominated the phenotype space and produced unstable sensitivity estimates. Sedation- and intubation-related distortions render the Glasgow Coma Scale highly vulnerable to systematic measurement error across different clinical centers. The five-system framework delivered a balanced exposure structure that preserved clinically relevant organ dysfunction data. Lambden et al. established that the neurological SOFA component remains one of the least stable metrics for multicenter critical care research [21].

The study has limitations. First, this was a retrospective analysis of derived public databases and residual confounding is unavoidable. Second, AF ascertainment was more granular in MIMIC-IV than in eICU, so the validation endpoint should be regarded as a pragmatic external validation endpoint rather than adjudicated incident AF. Third, the reduced validation model excluded some derivation variables because they were not transportable across datasets or had excessive validation missingness. Fourth, complete-case analysis was used because missingness in modeled variables was low in derivation and modest in validation, but the missing-data mechanism

was not formally tested. Fifth, neurological SOFA was excluded from the primary phenotype, so this study does not claim to represent the full six-component SOFA space in the main analysis.

Despite these limitations, the study has practical implications. The leading-organ phenotype can be derived early, is rule based, and does not require complex model explanation methods. Cardiovascular-leading and respiratory-leading trajectories may identify patients in whom closer rhythm surveillance, trigger modification, and targeted preventive strategies warrant further prospective study [8–10, 12, 20].

Conclusions

A five-system leading-organ phenotype derived from the first 72 hours of prolonged ICU admission identified distinct risk trajectories for subsequent atrial fibrillation. The cardiovascular-leading phenotype maintained the most consistent association with later NOAF across derivation and external validation cohorts. This temporally ordered organ-dysfunction framework complements existing critical-care risk models by identifying early arrhythmogenic states..

Declarations

Ethics approval and consent to participate: The source databases were approved by the Institutional Review Boards of the Massachusetts Institute of Technology and Beth Israel Deaconess Medical Center, and informed consent was waived because all protected health information had been removed. The eICU-CRD is a deidentified public multicenter database. The study plan was additionally approved by the Academic Board of the Department of Cardiology, Bezmialem Vakif University Faculty of Medicine (meeting no. 3, decision no. 12, dated 12 March 2026). The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication: Not applicable.

Data availability: The datasets analyzed during the current study are available in the PhysioNet repository: MIMIC-IV (version 3.1), https://doi.org/10.13026/kpb9-mt58; and eICU Collaborative Research Database (version 2.0), https://doi.org/10.13026/C2WM1R. Access to both datasets is provided to credentialed users who complete the required training and sign the applicable data use agreement through PhysioNet.

Funding: None.

Declaration of Competing Interest: The authors declare no competing interests.

Author contributions: H.B.I. conceived the study, designed the analysis, and supervised the project. H.B.I. and S.T.Y. performed data curation and statistical analysis. H.B.I. drafted the manuscript. S.T.Y., S.B.,E.S., E.O. and F.K. contributed to methodology development and interpretation of results. S.B. and F.K. prepared figures and tables. C.A. provided clinical oversight and critical revision of the manuscript. All authors reviewed, edited, and approved the final manuscript.

Use of generative AI: Artificial intelligence-assisted technologies were consulted to optimize manuscript readability. The authors independently verified all text, data, and scientific conclusions prior to submission.

Acknowledgements: None.

References

1. Vincent, J. L. et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med.* **22** (7), 707–710. 10.1007/BF01709751 (1996).
2. Johnson, A. E. W. et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci. Data*. **10** (1), 1. 10.1038/s41597-022-01899-x (2023).
3. Pollard, T. J. et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci. Data*. **5**, 180178. 10.1038/sdata.2018.178 (2018).
4. von Elm, E. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann. Intern. Med.* **147** (8), 573–577. 10.7326/0003-4819-147-8-200710160-00010 (2007).
5. Bosch, N. A., Cimini, J. & Walkey, A. J. Atrial fibrillation in the ICU. *Chest* **154** (6), 1424–1434. 10.1016/j.chest.2018.03.040 (2018).
6. Bedford, J. P. et al. New-onset atrial fibrillation in intensive care: epidemiology and outcomes. *Eur. Heart J. Acute Cardiovasc. Care*. **11** (8), 620–628. 10.1093/ehjacc/zuac080 (2022).
7. O'Bryan, L. J. et al. Managing new-onset atrial fibrillation in critically ill patients: a systematic narrative review. *BMJ Open.* **10** (3), e034774. 10.1136/bmjopen-2019-034774 (2020).
8. Bedford, J. P., Ede, J. & Watkinson, P. J. Triggers for new-onset atrial fibrillation in critically ill patients. *Intensive Crit. Care Nurs.* **67**, 103114. 10.1016/j.iccn.2021.103114 (2021).
9. Bedford, J. P. et al. Risk factors for new-onset atrial fibrillation during critical illness: a Delphi study. *J. Intensive Care Soc.* **23** (4), 414–424. 10.1177/17511437211022132 (2022).
10. Bosch, N. A., Cohen, D. M. & Walkey, A. J. Risk factors for new-onset atrial fibrillation in patients with sepsis: a systematic review and meta-analysis. *Crit. Care Med.* **47** (2), 280–287. 10.1097/CCM.0000000000003560 (2019).
11. Corica, B., Romiti, G. F., Basili, S. & Proietti, M. Prevalence of new-onset atrial fibrillation and associated outcomes in patients with sepsis: a systematic review and meta-analysis. *J. Pers. Med.* **12** (4), 547. 10.3390/jpm12040547 (2022).
12. Wetterslev, M. et al. Treatments of new-onset atrial fibrillation in critically ill patients: a systematic review with meta-analysis. *Acta Anaesthesiol. Scand.* **66** (4), 432–446. 10.1111/aas.14032 (2022).
13. Xu, Z. et al. Sepsis subphenotyping based on organ dysfunction trajectory. *Crit. Care*. **26**, 197. 10.1186/s13054-022-04071-4 (2022).
14. Guan, C. et al. Interpretable machine learning model for new-onset atrial fibrillation prediction in critically ill patients: a multi-center study. *Crit. Care*. **28**, 349. 10.1186/s13054-024-05138-0 (2024).
15. Wen, S. L. et al. The association between atrial fibrillation burden and all-cause mortality in critically ill survivors: a retrospective study from the MIMIC-IV database. *Nurs. Crit. Care*. **31** (1), e70334. 10.1111/nicc.70334 (2026).
16. Nguyen, V. & Mittal, R. Machine learning prediction of ICU mortality and length of stay in atrial fibrillation: a MIMIC-IV/MIMIC-III study. *Healthc. (Basel)*. **14** (3), 356. 10.3390/healthcare14030356 (2026).
17. Stevens, R. J. & Poppe, K. K. Validation of clinical prediction models: what does the calibration slope really measure? *J. Clin. Epidemiol.* **118**, 93–99. 10.1016/j.jclinepi.2019.09.016 (2020).
18. Wu, Y. et al. Serum electrolyte concentrations and risk of atrial fibrillation: an observational and Mendelian randomization study. *BMC Genom.* **25** (1), 280. 10.1186/s12864-024-10197-2 (2024).

19. Ho, K. M., Sheridan, D. J. & Paterson, T. Use of intravenous magnesium to treat acute onset atrial fibrillation: a meta-analysis. *Heart* **93** (11), 1433–1440. 10.1136/hrt.2006.111492 (2007).
20. Chyou, J. Y. et al. Atrial fibrillation occurring during acute hospitalization: a scientific statement from the American Heart Association. *Circulation* **147** (15), e676–e698. 10.1161/CIR.0000000000001133 (2023).
21. Lambden, S., Laterre, P. F., Levy, M. M. & Francois, B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. *Crit. Care*. **23**, 374. 10.1186/s13054-019-2663-7 (2019).
22. Shaver, C. M. et al. Atrial fibrillation is an independent predictor of mortality in critically ill patients. *Crit. Care Med.* **43** (10), 2104–2111. 10.1097/CCM.0000000000001166 (2015).
23. Moss, T. J. et al. New-onset atrial fibrillation in the critically ill. *Crit. Care Med.* **45** (5), 790–797. 10.1097/CCM.0000000000002325 (2017).

Figures

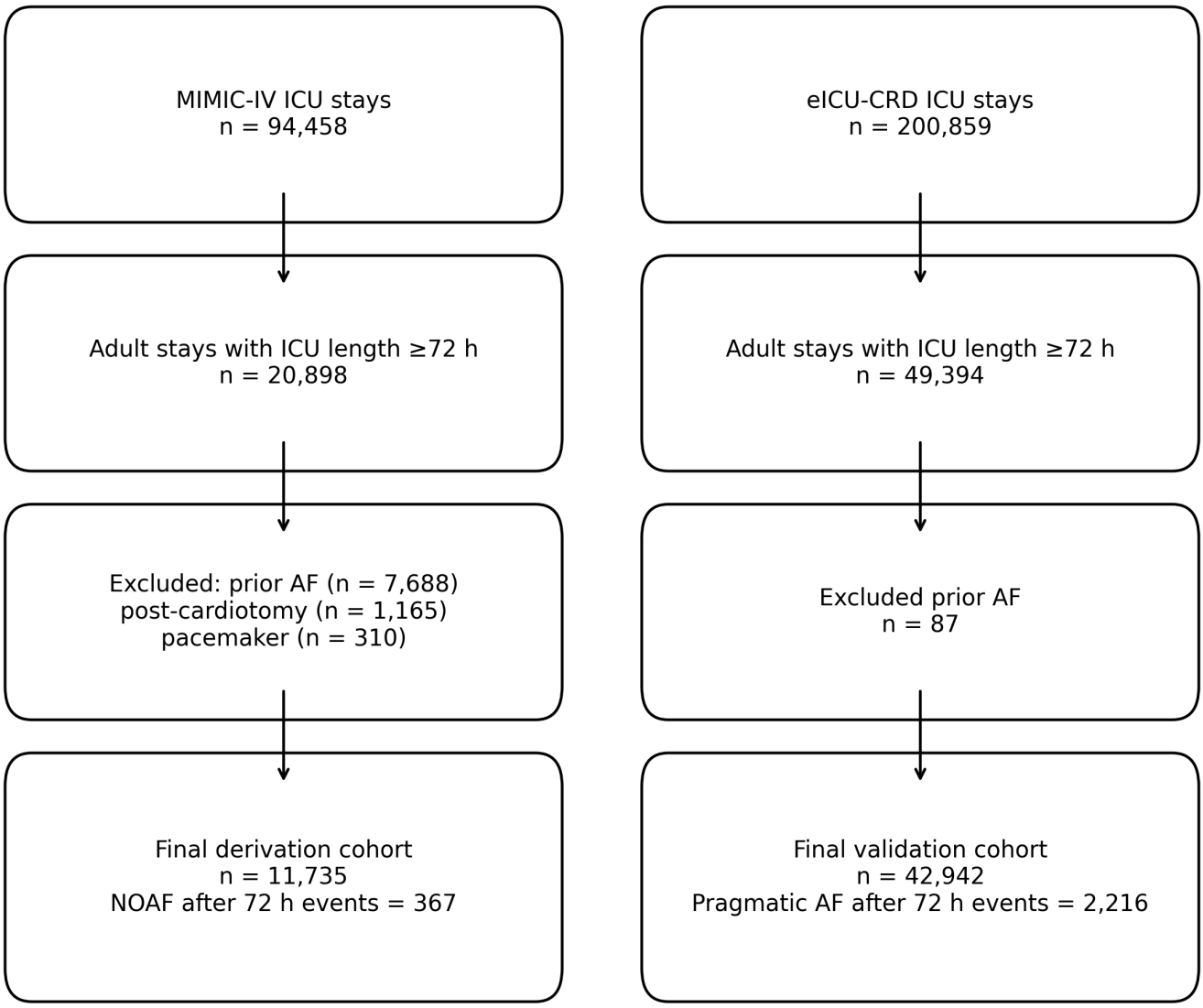


Figure 1

Study flow diagram for MIMIC-IV derivation and eICU external validation cohorts.

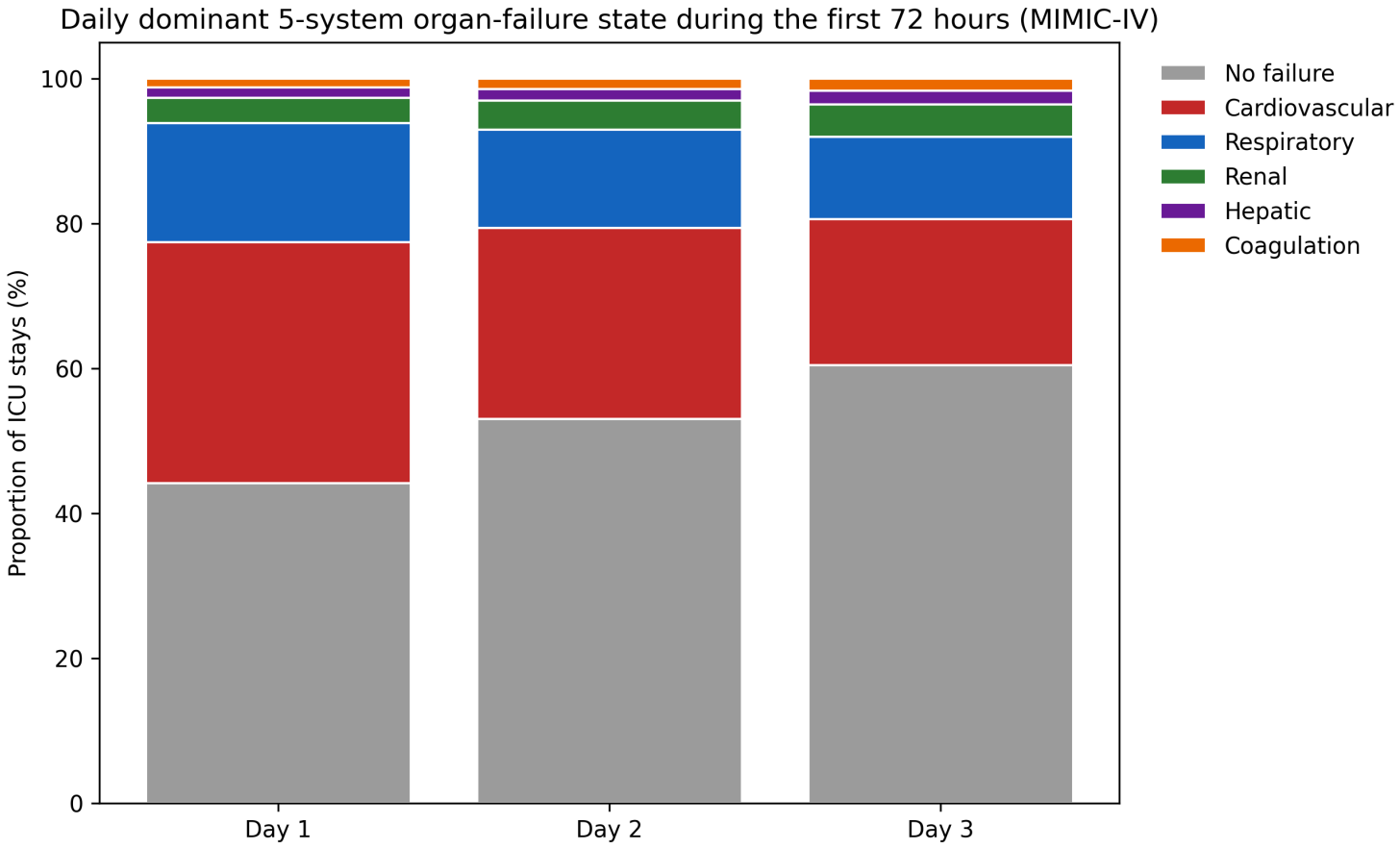


Figure 2

Daily dominant five-system organ-failure state during the first 72 hours in the MIMIC-IV derivation cohort.

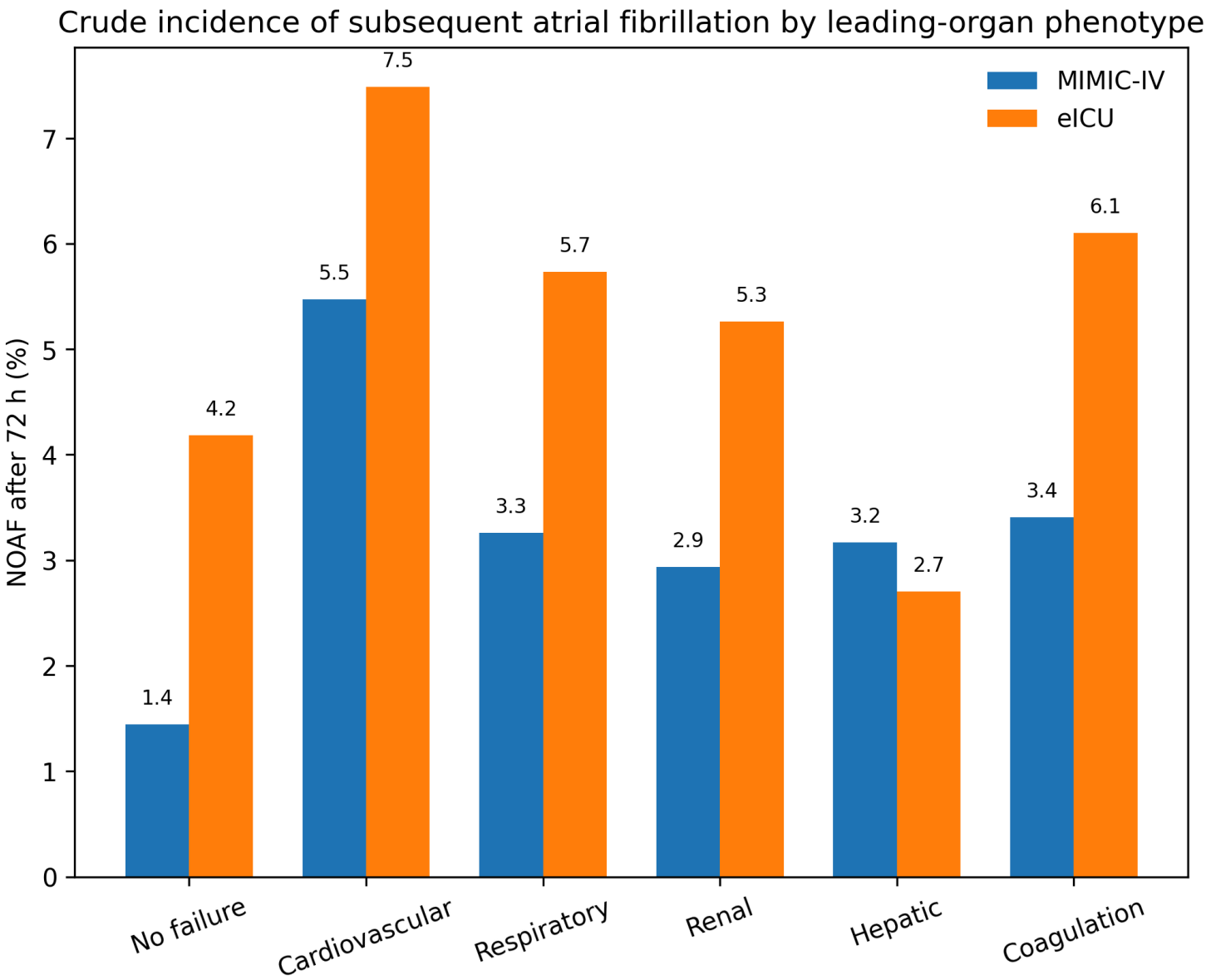


Figure 3

Crude incidence of subsequent atrial fibrillation by leading-organ phenotype in MIMIC-IV and eICU.

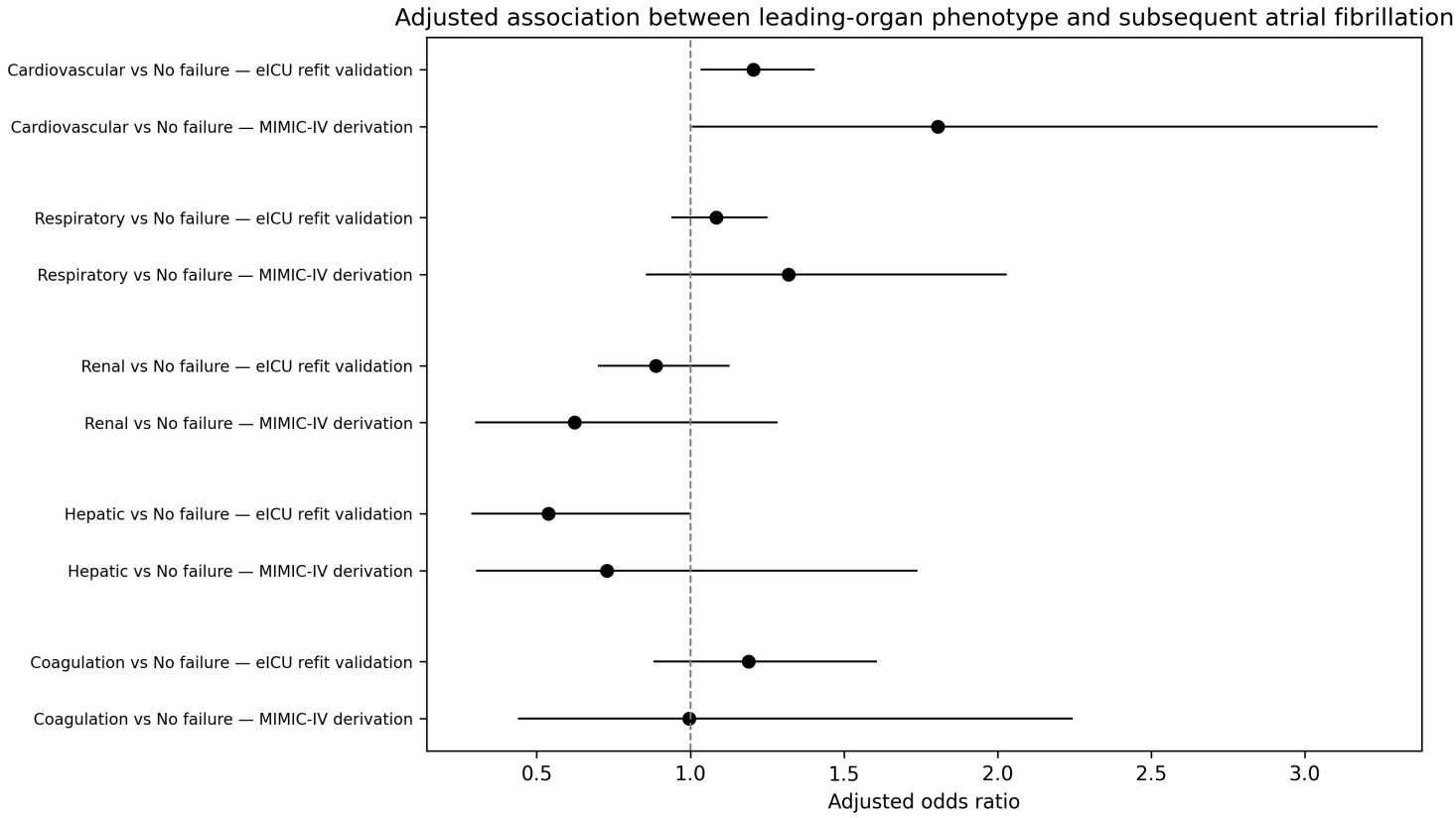


Figure 4

Adjusted association between leading-organ phenotype and subsequent atrial fibrillation in derivation and validation cohorts.

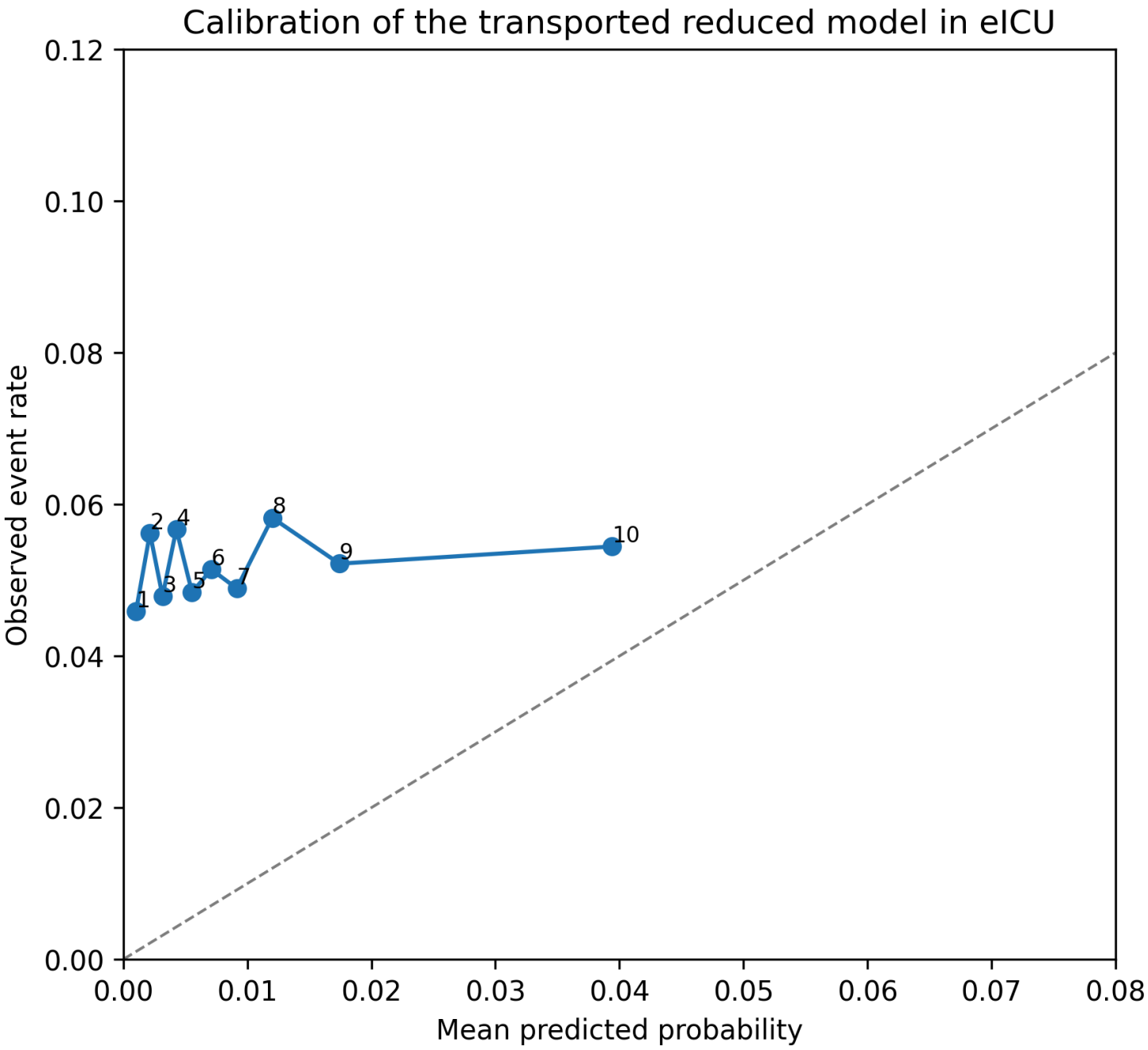


Figure 5

Calibration of the transported reduced derivation model in the eICU validation cohort.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 04SupplementaryInformation.docx