JAMIA Open, 2025, **8(1)**, ooae158
https://doi.org/10.1093/jamiaopen/ooae158
Research and Applications

Research and Applications

Multi-modal prediction of extracorporeal support—a resource intensive therapy, utilizing a large national database

Daoyi Zhu, BS[1], Bing Xue, PhD[1], Neel Shah, MD[2,3], Philip Richard Orrin Payne, PhD[4,5], Chenyang Lu, PhD[1,2], Ahmed Sameh Said, MD, PhD*[,2,3,5]

[1]Department of Computer Science and Engineering, Washington University in St Louis, St Louis, MO 63130, United States, [2]Artificial Intelligence (AI) for Health Institute (AIHealth), Washington University in St Louis, St Louis, MO 63130, United States, [3]Department of Pediatrics, Washington University in St Louis, St Louis, MO 63110, United States, [4]Department of Medicine, Washington University in St Louis, St Louis, MO 63110, United States, [5]Institute of Informatics, Data Science and Biostatistics, Washington University in St Louis, St Louis, MO 63110, United States

*Corresponding author: Ahmed Sameh Said, MD, PhD, Pediatric Critical Care, Washington University in St Louis, 660 South Euclid Avenue, MSC 8208-16-01, St Louis, MO 63110, United States (said_a@wustl.edu)

Abstract

Objective: Extracorporeal membrane oxygenation (ECMO) is among the most resource-intensive therapies in critical care. The COVID-19 pandemic highlighted the lack of ECMO resource allocation tools. We aimed to develop a continuous ECMO risk prediction model to enhance patient triage and resource allocation.

Material and Methods: We leveraged multimodal data from the National COVID Cohort Collaborative (N3C) to develop a hierarchical deep learning model, labeled "PreEMPT-ECMO" (Prediction, Early Monitoring, and Proactive Triage for ECMO) which integrates static and multi-granularity time series features to generate continuous predictions of ECMO utilization. Model performance was assessed across time points ranging from 0 to 96 hours prior to ECMO initiation, using both accuracy and precision metrics.

Results: Between January 2020 and May 2023, 101 400 patients were included, with 1298 (1.28%) supported on ECMO. PreEMPT-ECMO outperformed established predictive models, including Logistic Regression, Support Vector Machine, Random Forest, and Extreme Gradient Boosting Tree, in both accuracy and precision at all time points. Model interpretation analysis also highlighted variations in feature contributions through each patient's clinical course.

Discussion and Conclusions: We developed a hierarchical model for continuous ECMO use prediction, utilizing a large multicenter dataset incorporating both static and time series variables of various granularities. This novel approach reflects the nuanced decision-making process inherent in ECMO initiation and has the potential to be used as an early alert tool to guide patient triage and ECMO resource allocation. Future directions include prospective validation and generalizability on non-COVID-19 refractory respiratory failure, aiming to improve patient outcomes.

Lay Summary

Extracorporeal membrane oxygenation (ECMO) is a life-saving therapy used in intensive care units (ICU). It requires the most ICU resources, and using ECMO is the most complex decision made in the ICU. The COVID-19 pandemic highlighted significant challenges with allocating ECMO resources. In response, we developed PreEMPT-ECMO (Prediction, Early Monitoring, and Proactive Triage for ECMO), a machine learning model to predict ECMO use in ICUs. The model was developed using data from over 100 000 patients in the National COVID Cohort Collaborative (N3C) between 2020 and 2023. The data included static variables (age, medical history) and time-series variables (changes in laboratory results and vital signs). PreEMPT-ECMO generated continuous prediction of ECMO risk up to 96 hours in advance. The model outperformed traditional methods in accuracy and precision, providing more reliable predictions of ECMO use. This methodology mirrors the complex decision-making clinicians face when deciding to use ECMO. Identifying patients early helps clinicians prioritize resources for those most at risk or transfer them to ECMO-experienced centers before they become too sick. Future work will test this model in broader ICU populations to expand its use beyond COVID-19 and measure its impact on outcomes of critically ill patients.

Key words: ECMO; COVID-19; machine learning; prediction; early alert; resource allocation.

Introduction

Extracorporeal membrane oxygenation (ECMO) is a critical life-saving therapy used in intensive care units (ICUs) for patients with severe refractory respiratory or cardiac failure.[1–3] ECMO is considered the most resource-intensive ICU therapy, posing significant challenges, especially during healthcare resource constraints. The decision to initiate ECMO is arguably one of the most complex in the ICU, requiring the integration of a wide range of patient-specific variables, including their dynamic physiological responses to therapeutic interventions. The severe acute respiratory syndrome coronavirus-2 (COVID-19) pandemic underscored these challenges. This emphasized the urgent need for advanced predictive models to identify patients at highest risk of requiring ECMO support and to

Received: November 18, 2024; **Revised:** December 13, 2024; **Editorial Decision:** December 20, 2024; **Accepted:** December 24, 2024

ensure the precise and timely deployment of ECMO and other resource-intensive therapies, in order to rescue patients.[4–8] During the pandemic's peak, ECMO use was often reserved for the most experienced centers, further stressing the necessity for efficient resource management in already overwhelmed healthcare systems and early patient identification and transfer.[4,9–11]

Existing predictive models, such as the ForecastECMO developed by Xue et al., attempted to address these challenges by predicting ECMO use in advance.[12] However, ForecastECMO had notable limitations: it was derived from a single-center dataset, utilized ensemble gradient boosting tree (GBT) models to predict at fixed 2-hour intervals from ICU admission, and did not incorporate the temporal evolution of patient data—a crucial factor for high-risk decision-making in clinical settings. These limitations emphasize the need for more sophisticated models capable of continuous prediction and integration of time series data. There is growing interest in the use of multi-modal, multi-institutional datasets to validate whether machine learning (ML) can provide consistent, clinically relevant guidance with sufficient lead time.[13–15]

The National COVID Cohort Collaborative (N3C) was established to meet the urgent need for comprehensive, multi-institutional data during the pandemic.[16] This collaboration aggregated data from numerous centers across the United States, creating a robust platform for developing and validating predictive models. The N3C dataset includes extensive demographic, comorbidity, and time series data, offering a rich resource for a more nuanced and generalizable approach to developing prognostic and therapeutic models for therapies like ECMO. Integrating such a diverse dataset aims to overcome the limitations of single-center studies and enhance the reliability and validity of predictive models in varied clinical settings. Prior research using the N3C database has demonstrated that ML models could accurately predict clinical severity from clinical data within the first 24 hours of admission.[17] Additionally, ML methodologies have been used to develop ECMO predictive models focused on survival, using international registry data.[18]

In response to these challenges, we aimed to develop a predictive model for early identification of ECMO use in critically ill COVID-19 patients using N3C data. Our primary objective was to validate and adapt a model leveraging multicenter data to provide a continuously predicting framework capable of handling evolving time series features. Furthermore, we sought to evaluate this model across several time horizons and compare our hierarchical model to widely used ML models.

Methods

Setting and data sources

This study was approved by the Washington University in St Louis Institutional Review Board (#202011004, approved on 11/02/2020). Data were accessed through the N3C Data Enclave platform (https://covid.cd2h.org/enclave/). A study protocol approved by the N3C Data Access Committee is housed in the N3C Enclave. This secure platform provides access to harmonized clinical data from contributing centers on the clinical characteristics and outcomes of patients tested for or diagnosed with COVID-19. To date, this includes data from over 22 million patients, over 8 million confirmed COVID-19 positive cases across 84 sites and with over 33 billion rows of data. We included demographics, treatment, and flowsheet data on patients confirmed to have COVID-19 by polymerase chain reaction (PCR) testing from January 2020 to May 2023. Inclusion criteria included admission to an ICU for at least 24 hours to ensure sufficient data for model development. Similar to our prior ECMO utilization prediction model, we excluded patients less than 3 years old, over 70 years old and with a calculated body mass index (BMI) over 45 kg/m^2 due to the controversies regarding ECMO use in these patients during the pandemic.[12] The primary outcome was provision of ECMO support and time in hours from ICU admission to ECMO initiation. We reported the study results using the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD+AI) reporting guideline[19] (Digital Content S1).

Data and data processing

Available features were categorized as static variables available at admission and time series variables collected during the ICU stay. Static features included patient demographics and comorbidities (eg, age, BMI, and Charlson Comorbidities Index). The time series variables included vital signs, therapeutics, and laboratory values such as heart rate, blood pressure, blood glucose and hemoglobin levels, and supplemental oxygen flow rates, amongst others, with average reporting frequencies ranging between 1 and 20 hours.

Missing data and feature engineering

To account for variable missing rates between patients and reporting institutions, time series features were selected in a 2-step approach. First, the top 100 features by availability were identified. Next, the features most relevant for ECMO decision making were identified (by authors A.S. and N.S.), resulting in 47 variables. To address the heterogeneity in time series data caused by varying observation frequencies (granularity) and missing rates, we adjusted the sampling rate of each variable to maintain sufficient sequence length and keep the missing rate below 60%. Consequently, the included time series variables were grouped into 3 granularity groups with sampling rates of 4 , 16, and 24 hours respectively.

ML approach

Building on previous efforts in developing predictive models to incorporate time series variables,[13,20–22] we developed a novel predictive model, PreEMPT-ECMO (Prediction, Early Monitoring, and Proactive Triage for ECMO). This model employs a hierarchical structure with 3 modules: a static network module to capture insights from the static features; a multi-granular time series network module that learns the 3 levels of grouped time series features; and an overarching network that integrates the representations from both static and time series networks to generate predictions, as depicted in Figure 1. As illustrated in the figure, the time series network modules were initiated by the representations of the static network module based on the patients' static features. We used a Long Short-Term Memory (LSTM) network as the foundational architecture for each time series granularity, adjusting input sequence lengths to capture variations in feature availability temporal scales. The overarching network combined the representations from both the static network module and the multi-granular time series network module to

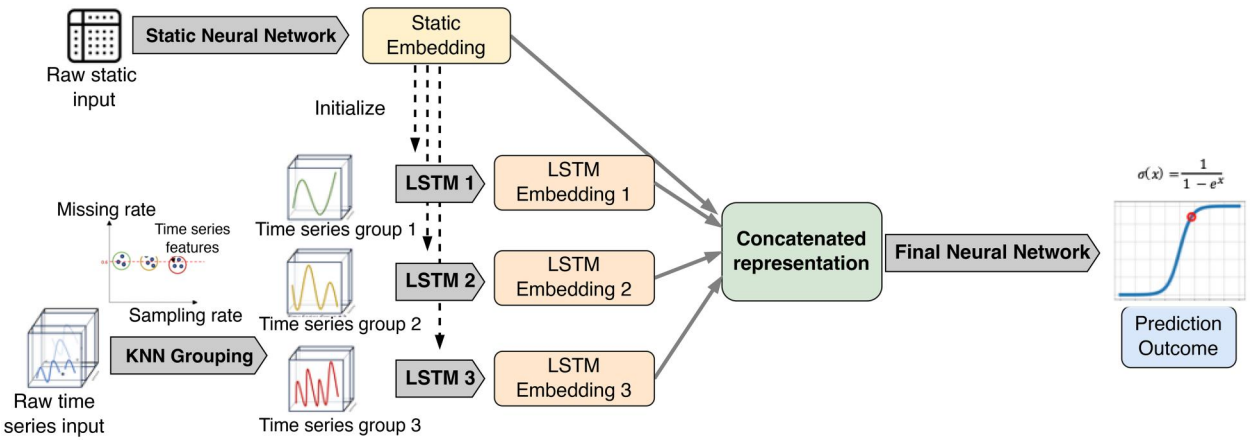

Figure 1. PreEMPT-ECMO model architecture. Schematic diagram of the PreEMPT-ECMO model architecture. The static features were fed into a static network module to extract demographics and comorbidities representations. The raw time series feature inputs were grouped into 3 levels of granularity groups based on sampling frequency and missingness. The static representations were used to initiate each LSTM backbone and concatenated with the sequential layers' outputs into the final model to learn the predicted probability of ECMO utilization. The final model output provided the probability and timing of ECMO initiation. LSTM, long short term memory; ECMO, extracorporeal membrane oxygenation.

predict the probability of ECMO use and the time from ICU admission to ECMO initiation.

To develop a comprehensive assessment of model performance in the face of anticipated low ECMO incidence rate, we performed random 5-fold cross-validation shuffling. Each iteration used a different stratified fold for model evaluation, and the remaining folds for model training. To improve model generalizability for the ECMO cases, each ECMO patient was up sampled during training stage by predicting at various timepoints, introducing a random time interval t before ECMO initiation. Consequently, the ML model only took the time series observations collected t hours before ECMO initiation as inputs.

Model performance and evaluation

To evaluate the model performance, widely used linear and nonlinear baseline ML models were developed for comparison. Linear models included support vector machine (SVM) and logistic regression (LR), and nonlinear models included random forest (RF) and extreme gradient boosting tree (XGBoost). The SVM, LR, and RF models were implemented using the Python Sklearn package and the XGBoost model using the XGBoost package. Since most baseline models are not designed to handle multi-modal inputs (combining static and time series features), we extracted the first and second-order statistical features to characterize each time series distribution, including metrics as maximum, mean, minimum, kurtosis, skewness, and crossing rates. All code and model configurations for this study are accessible on GitHub, providing the necessary resources to reproduce and extend the analyses conducted.

During model evaluation, 2 performance measures were recorded at every t in each iteration: the area under the receiver operating characteristic curve (AUROC), and the area under the precision recall curve (AUPRC). To compare ECMO and non-ECMO patients at each time point t, we aligned non-ECMO patients with ECMO patients based on similar elapsed time since ICU admission. This alignment reflects the clinical scenario, where predictions are more challenging at larger time points (larger t) as both patient groups are newly admitted to the ICU. In contrast, predictions are easier at smaller time points (smaller t) since ECMO patients typically deteriorate from hospital admission to ECMO initiation, while non-ECMO patients generally improve until discharge. Feature contribution to model performance was evaluated using SHapley Additive exPlanation (SHAP) analysis.[23,24] This model-agnostic analysis provides an illustration of the contribution of various features in models' performance in both a generalized approach and on specific subgroup levels.

Data are presented as median and interquartile range (IQR) for quantitative variables and number and percentage for qualitative variables, unless otherwise specified. Wilcoxon rank sum test or Chi square tests were used for group comparisons. A 2-sided P-value $\leq .05$ was considered statistically significant.

Results

During the study period, we identified a total of 19 027 360 patients in the N3C cohort. After implementation of the inclusion and exclusion criteria 101 400 patients were included in the model development and validation, with 1298 (1.28%) supported on ECMO (Figure 2).

Cohort characteristics

Most of the patients were female (58.9%) with the majority being of non-Caucasian race (55.5%) with a median and IQR age of 50.46 [39, 64] years, weight of 86.38 [70.26, 108.65] kg and height of 168.04 [160.66, 177.47] cm. The overall mortality rate was 13.4% with a median ICU length of stay (LOS) of 12 [6, 24] days. The ECMO cohort were predominantly male (55.1%) as compared to the non-ECMO patients (40.9%), $P < .0001$. As anticipated, there was a higher mortality rate and thus shorter ICU LOS in the ECMO cohort (40%) of 10 [4, 19] days as compared to the non-ECMO cohort (13%) of 12,[6,24] $P < .0001$. Table 1 shows the overall patient characteristics of the included cohort and the comparison between the ECMO and non-ECMO cohorts.

Model performance

To evaluate model's performance at different time horizons, we calculated each model's AUROC and AUPRC score from

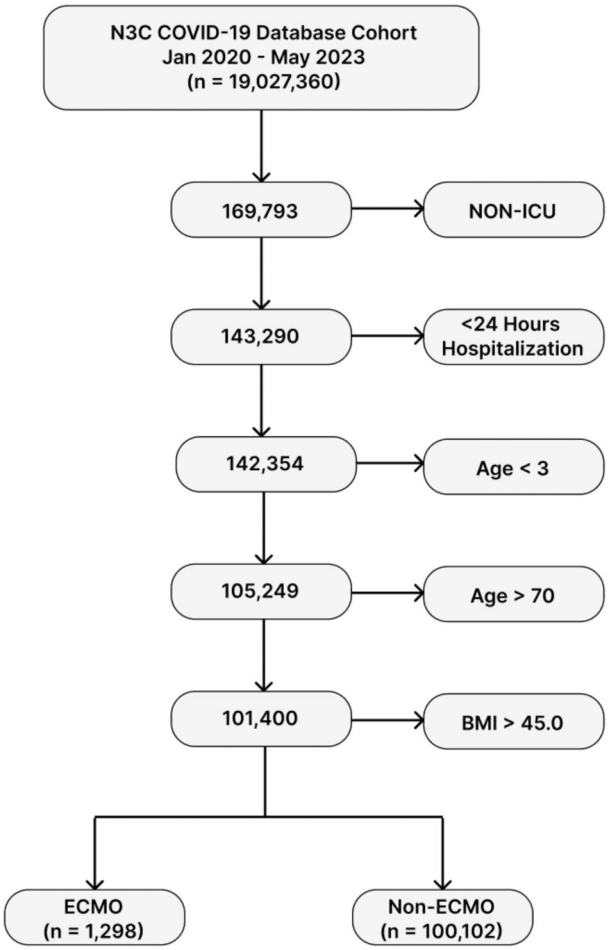

Figure 2. Patient selection diagram. A flowchart of the patient selection process. A total of 19 027 360 patients were identified from the N3C data enclave between January 2020 and May 2023, with 169 793 patients admitted to an ICU and 143 290 with ICU length of stay more than 24 hours. Patients under the age of 3 were excluded as their normal ranges for vital signs differ markedly from those of adult patients, who make up the bulk of individuals supported on ECMO. Finally, we excluded patients with BMI larger than 45 kg/m^2 and/or age greater than 70 years old following the institutional ECMO exclusion criteria. The final cohort included 101 400 patients with 1298 supported on ECMO. N3C, National COVID Cohort Collaborative; ICU, intensive care unit; BMI, body mass index; ECMO, extracorporeal membrane oxygenation.

$t = 0$ to $t = 96$ hours prior to ECMO initiation. The proposed model, PreEMPT-ECMO, outperformed LR, SVM, RF, XGBoost in terms of accuracy as represented by AUROC (Figure 3A) across all tested time horizons and in terms of precision as measured by AUPRC at most time horizons before $t = 80$ (Figure 3B).

To focus on potentially clinically actionable time points that could influence patient triage or ECMO resource allocation, we evaluated the model performance at 12, 24, and 48 hours prior to ECMO initiation. PreEMPT-ECMO demonstrated superior performance to comparative models in both accuracy (AUROC) and precision (AUPRC). Table 2 presents the accuracy and precision metrics for all the compared models at these key time points. At each time point, 12, 24, and 48 hours prior to ECMO initiation, PreEMPT-ECMO had significantly higher AUROC and AUPRC with P values ranging between $<.05$ and $<.001$ as compared to LR, SVM, and RF models. PreEMPT-ECMO also outperformed the comparator XGBoost model in both performance metrics at each time points except for AUPRC at 48 hours, where statistical significance was not achieved.

Feature importance

To evaluate the most important features in the PreEMPT-ECMO model performance, we performed a SHAP analysis. For the overall model performance, hemoglobin levels, respiratory rate, serum potassium and calcium levels and hematocrit were the most salient features (Figure 4A). These features are recognized to be of clinical importance in gauging the severity of both respiratory failure, systemic oxygenation, and end organ function.[25–28] When repeating the SHAP analysis for only the patients predicted to receive ECMO support, again hemoglobin, hematocrit, and serum potassium were identified in the top 5 features of importance. Additionally, measured pH of arterial blood and fraction of inspired oxygen were identified to be significant (Figure 4B). The identification of both pH of arterial blood and fraction of inspired oxygen highlight the model's incorporation of the most severe forms of respiratory failure and the impact of escalating therapeutics to the prediction of ECMO use.[29–31] For the SHAP analysis of the model performance in identifying the patients not supported on ECMO, hemoglobin, respiratory rate, serum potassium and calcium, and hematocrit were identified as the features of highest contribution to the model performance (Figure 4C). The similarity in identified feature importance to those in the overall model performance could be reflective of the significant class imbalance between ECMO and non-ECMO patients in the cohort.

Individual feature importance

To understand differences in peak predictive performance and feature importance between patients, we evaluated the change in model performance and feature contribution through SHAP analysis across time horizons. The SHAP analysis was performed at the pivot point when the prediction probability surged. For different patients, the most salient features at the pivot points varied. In addition, the importance of these features at the pivot points had varying contributions across time horizons, further emphasizing the model's adaptability. Figure 5 illustrates 3 examples of the model predictive performance demonstrating the differences in feature contribution importance and change over time horizons.

Discussion

We present PreEMPT-ECMO, a hierarchical ML model designed to predict ECMO use continuously utilizing a large national COVID-19 database encompassing 3 years and over 100 000 patients. This model achieved high accuracy and precision across several prediction horizons following ICU admission. PreEMPT-ECMO outperformed traditional LR, SVM, RF, and XGBoost classifiers in predicting ECMO utilization. Moreover, the model exhibited strong clinical explainability and credibility through SHAP analysis. Case studies on individual feature importance further highlighted the model's adaptability and individualized predictive ability, revealing how certain factors shifted in importance throughout the prediction windows. This feature enables the examination of ECMO use probabilities at different time intervals

Table 1. Characteristics of studied cohort.

Median, [IQR]	Entire cohort ($n = 101\,400$)	ECMO ($n = 1298$)	Non-ECMO ($n = 100\,102$)	*P* value
Male, (%)	41 628 (41.1%)	715 (55.1%)	40 913 (40.9%)	<.0001
Caucasian, (%)	45 162 (44.5%)	554 (42.6%)	44 608 (44.6%)	.4125
Age, years	50.49 [39, 64]	48.20 [40, 59]	50.49 [39, 64]	<.005
Height, cm	168.04 [160.66, 177.47]	170.31 [162.61, 177.8]	168.01 [160.6,177.44]	.0002
Weight, kg	86.38 [70.26, 108.65]	84.03 [66.60, 101.06]	86.39 [70.31, 108.81]	<.0001
BMI, kg/m^2	30.06 [25.76, 35.06]	31.0 [27.10, 35.42]	30.04 [25.74, 35.06]	.006
ICU LOS, days	12 [6, 24]	10 [4, 19]	12 [6, 24]	<.0001
CCI	1 [0, 2]	1 [0, 2]	1 [0, 2]	<.0001
Mortality, (%)	13 637 (13.4%)	519 (40.0%)	13 118 (13.0%)	<.0001
Dementia	3386(3.3%)	1 (0.1%)	3385 (3.4%)	<.0001
PVD	4274(4.2%)	32 (2.5%)	4242 (4.3%)	.0020
COPD	4590(4.5%)	30 (2.3%)	4560 (4.5%)	.0001
Old MI	5806(5.7%)	37 (2.8%)	5769 (5.7%)	<.0001
CVD	842(0.8%)	6 (0.4%)	836 (0.8%)	.1878
CKD	10 833(10.6%)	121 (9.3%)	10 712 (10.7%)	.1205
CHF	11 705(11.5%)	121 (9.3%)	11 584 (11.5%)	.0132
AIDS	5 (0.0049%)	0 (0.0%)	5 (0.01%)	1
Hemiplegia	290(0.2%)	5 (0.3%)	285 (0.2%)	.6803
Tumor history	73 (0.07%)	0 (0.0%)	73 (0.07%)	.6509
Type 2 DM	20 670(20.3%)	11 (0.8%)	20 659 (20.6%)	<.0001
CNLD	298(0.3%)	5 (0.4%)	293 (0.3%)	.7236

P value by Chi square test for categorical features and Wilcoxon rank sum test for continuous variables.
Abbreviations: AIDS = acquired immune deficiency syndrome; BMI = body mass index; CCI = Charlson comorbidity index; CHF = congestive heart failure; CKD = chronic kidney disease; CNLD = chronic non-alcoholic liver disease; COPD = chronic obstructive pulmonary disease; CVD = cerebrovascular disease; DM = diabetes mellitus; ICU = intensive care unit; IQR = interquartile range; LOS = length of stay; MI = myocardial infarction; PVD = peripheral vascular disease.

and provides deeper insight into how individual variables impact the model's performance on each patient level.

These advancements mark significant improvements over the ForecastECMO model developed by Xue et al. While ForecastECMO relied on static and time series variables, the ensemble GBT model architecture could only handle the time series variables at fixed 2-hour prediction windows, risking the loss of important variables, and more importantly key trends occurring within these fixed intervals. In contrast, PreEMPT-ECMO built on prior work utilizing the multiscale data recurrent neural network approach, which more effectively captures the dynamic temporal changes in the time series variables. This approach allows for a nuanced and clinically applicable prediction across multiple time scales, making PreEMPT-ECMO particularly suited to ICU settings where patient conditions evolve rapidly. Moreover, this model included data from a large national cohort spanning multiple centers, enhancing its generalizability and applicability.

Throughout the COVID-19 pandemic, ECMO use continued to be a highly resource-intensive therapy, often reserved for the most critically ill patients. PreEMPT-ECMO shows promise as a valuable tool for early identification of patients at risk of receiving ECMO. While it does not supersede clinical judgment, it serves as an adjunct to predict the use of a high-resource therapy, allowing for more timely resource allocation or patient triage and transfer to ECMO capable centers to potentially improve outcomes on both individual patient and healthcare system levels. This approach aligns with the growing evidence supporting early referral and intervention for ECMO candidates, potentially reducing mortality and improving recovery rates.[10,32,33] Unlike other ML models focused on predicting mortality prior to and during ECMO, our model, like ForecastECMO, aims to identify patients at high-risk of receiving ECMO, well in advance of the imminent need, thereby providing clinicians a window for patient transfer or adjustments in care.

Although COVID-19 remains a public health concern, it no longer imposes the same level of strain on the global health system as during the peak of the pandemic. Nonetheless, our model introduces a novel hierarchical ML approach, combining granular temporal data with static variables to provide risk predictions at multiple time intervals with unique insights into predictive variables and their changes over time. We validated PreEMPT-ECMO using a large national database, even though ECMO deployment remains a relatively rare event. Future work may leverage similar models to predict the need for other rare but resource-intensive therapies, such as dialysis, organ transplantation, and specialized surgical interventions or the use of ECMO in non-COVID-19 cases. Additionally, advancing these models to handle data with even higher temporal resolution, such as minute-by-minute monitoring could further refine real-time prediction accuracy. Integrating continuous prediction frameworks would facilitate seamless integration into clinical workflows, enabling real-time alerts and updates. Multicenter prospective studies will be essential to validate these models in real-world settings, ensuring their reliability and effectiveness in diverse patient populations.

Despite its favorable performance in a large national COVID-19 database, our model has several limitations. The granularity and frequency of data grouping depended on availability and reporting frequency of the contributing institutions, but did not capture minute-to-minute variations, that could potentially enhance predictive precision. Additionally, although PreEMPT-ECMO demonstrated good generalizability and validity in a large national database, further validation in external, prospective studies is necessary. While PreEMPT-ECMO demonstrated high predictive performance for ECMO utilization, it is important to recognize the

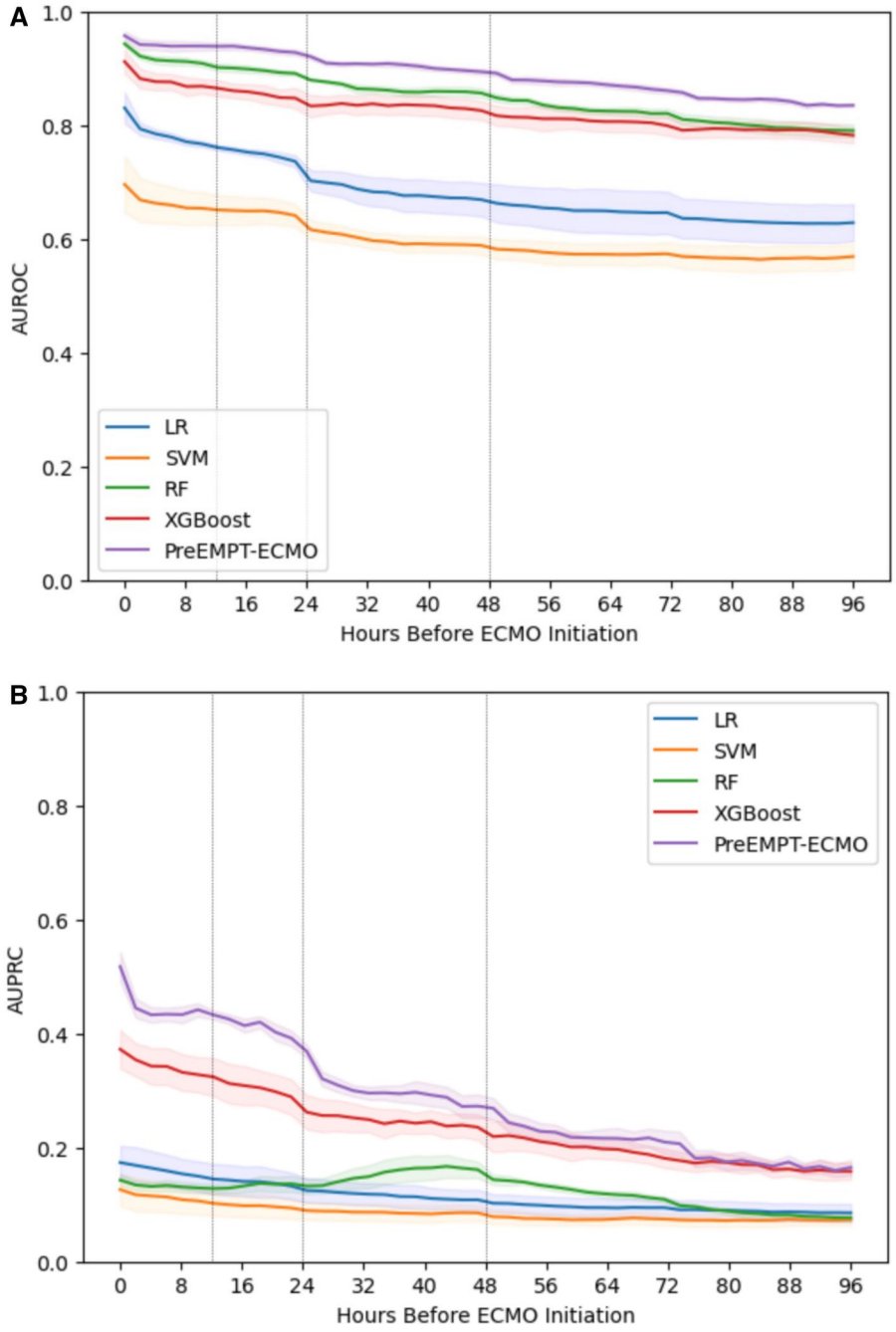

Figure 3. Comparative performance of PreEMPT-ECMO model across all studied timepoints. Comparison of the PreEMPT-ECMO model and other comparative models at predicting ECMO utilization at all studied timepoints. (A) AUROC at prediction horizons of 0-96 hours prior to ECMO initiation. PreEMPT-ECMO predicted ECMO utilization with higher accuracy than all comparator LR, SVM, RF, and GBT models at all time horizons. Potentially clinically important timepoints of 12, 24, and 48 hours where PreEMPT-ECMO outperformed all comparator models are delineated with vertical dotted lines. (B) AUPRC at prediction horizons of 0-96 hours prior to ECMO initiation PreEMPT-ECMO predicted ECMO utilization with higher precision than all comparator LR, RF, and GBT models at all time horizons. The PreEMPT-ECMO performed with higher precision than the SVM model in the 0-72 hours time horizons with comparable performance in the 72-96 hours timepoints. Potentially clinically important timepoints of 12, 24, and 48 hours where PreEMPT-ECMO outperformed all comparator models are delineated with vertical dotted lines. ECMO, extracorporeal membrane oxygenation; LR, logistics regression; SVM, support vector machine classifier; RF, random forest classifier; GBT, gradient boosting tree classifier; AUROC, area under receiver operator curve; AUPRC, area under precision recall curve.

inherent variability in ECMO indications and practices across centers. ECMO initiation is influenced by numerous factors, including disease etiology, patient risk stratification, and institutional thresholds for cannulation. Furthermore, the distinction between ECMO modalities—veno-venous or veno-arterial—each with differing indications and outcomes, adds to this heterogeneity. A more granular analysis of ECMO subtypes and their specific indications could enhance future model iterations. Although our study focused on predictive feasibility, it is worth noting that ECMO's survival

Table 2. Comparison of PreEMPT-ECMO model performance at select prediction time horizons prior to ECMO.

	Time point	PreEMPT-ECMO	LR	SVM	RF	GBT
AUROC	12 hours	0.94	0.76[b]	0.65[b]	0.9[a]	0.86[a]
	24 hours	0.92	0.70[b]	0.62[b]	0.88[b]	0.83[a]
	48 hours	0.89	0.66[b]	0.58[b]	0.85[a]	0.82[a]
AUPRC	12 hours	0.43	0.14[b]	0.10[b]	0.13[b]	0.32[a]
	24 hours	0.37	0.12[b]	0.09[b]	0.13[b]	0.26[a]
	48 hours	0.27	0.10[b]	0.08[b]	0.14[b]	0.22

[a] *P* value <.05, paired *t* test comparison between PreEMPT-ECMO and baseline models.
[b] *P* value <.01, paired *t* test comparison between PreEMPT-ECMO and baseline models.
Abbreviations: AUPRC = area under precision recall curve; AUROC = area under receiver operator curve; ECMO = extracorporeal membrane oxygenation; GBT = gradient boosting tree; LR = logistic regression; LSTM = long short term memory; RF = random forest; SVM = support vector machine.

benefit remains complex and context dependent. Prior studies in pediatric populations suggest that ECMO initiation in lower-risk patients may paradoxically increase mortality, while providing clear survival benefit in higher-risk patients.[34,35] While these findings are not directly addressed in our analysis, they underscore the need to develop models that identify the patients most likely to benefit from resource-intensive therapies like ECMO.

The COVID-19 pandemic underscored the importance of robust and adaptive tools for managing high-risk, resource-intensive therapies. The development of PreEMPT-ECMO is a step towards addressing this need, however, ongoing model improvements-such as incorporating more granular data and validating performance across different healthcare settings-are essential. The integration of such models into electronic health records (EHRs) could provide real-time decision support to clinicians, enhancing patient care, and resource management during critical situations.

Conclusion

This study presents PreEMPT-ECMO, a hierarchical ML model capable of continuously predicting ECMO use—a high-resource-intensive therapy—with high accuracy across various time horizons within a large multicenter national COVID-19 database. By uniquely integrating granular time series data with static variables, this model provides valuable insights into feature importance over varying time periods, enhancing its clinical interpretability. Although these developments represent significant advancements, further refinement and validation are required. The ultimate goal is to develop robust, reliable tools that support clinicians in providing optimal care for critically ill patients, particularly amid unprecedent healthcare challenges such as the COVID-19 pandemic. The continued evolution of these tools will be crucial in ensuring they remain effective and relevant in an ever changing healthcare landscape.

Acknowledgments

The analyses described in this publication were conducted with data or tools accessed through the NCATS N3C Data Enclave https://covid.cd2h.org and N3C Attribution and Publication Policy v 1.2-2020-08-25b supported by NCATS Contract No. 75N95023D00001. This research was possible because of the patients whose information is included within the data and the organizations (https://ncats.nih.gov/n3c/resources/data-contribution/data-transfer-agreement-signatories) and scientists who have contributed to the on-going development of this community resource (https://doi.org/10.1093/jamia/ocaa196).

Author contributions

Daoyi Zhu and Bing Xue performed the data cleaning, analysis, model building, and validation; Neel Shah and Ahmed Sameh Said provided conceptualization, clinical interpretation, manuscript writing; Chenyang Lu, and Philip Richard Orrin Payne provided guidance on data analysis, model building, and validation. All authors participated in manuscript revision and editing.

Supplementary material

Supplementary material is available at *JAMIA Open* online.

Funding

A.S.S. has received research support from the Children's Discovery Institute Faculty Development Award at Washington University in St Louis and St Louis Children's Hospital. For the remaining authors none were declared.

Conflicts of interest

None of the authors have any conflicts of interest to declare that affected the conduct or presentation of this research.

Data availability

The data underlying this article are available through the Data Use Request process at the N3C collaborative and approval by the Data Access Committee.

Tweet

Predictive models of ECMO use in COVID-19 are possible with good accuracy and precision, guiding resource allocation in states of system-wide stress.

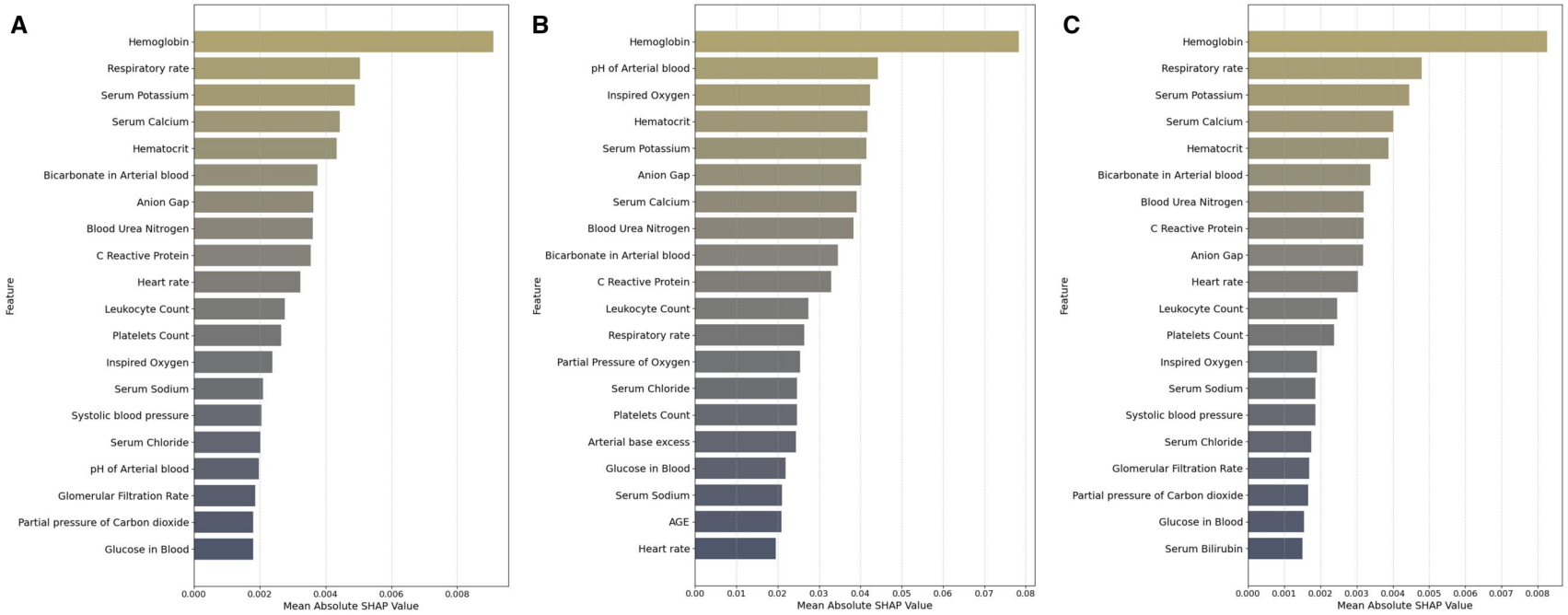

Figure 4. Feature contribution in PreEMPT-ECMO model performance by SHAP analysis. (A) Analysis of feature contribution by SHAP analysis for the total patient cohort shows that the most contributing features were measured hemoglobin, respiratory rate, serum potassium, serum calcium and hematocrit, all features of clinical significance in the quantification of COVID-19 associated respiratory failure, hypoxemia, and end organ function that impact the ECMO decision making. (B) Analysis of feature contribution by SHAP analysis for the patients predicted to be supported on ECMO, showing the most contributing features to be measured hemoglobin, pH of arterial blood, fraction of supplemental oxygen, hematocrit, and serum potassium levels. The addition of both pH of arterial blood and fraction of supplemental oxygen for the positive cohort, highlights the model's inclusion of features representing the progression of COVID-19 associated respiratory failure and the accompanying escalation in therapeutics. (C) Analysis of feature contribution by SHAP analysis for the patients not predicted to receive ECMO support. The most relevant features included measured hemoglobin, respiratory rate, serum potassium, serum calcium, and hematocrit. The similarity to features deemed to be most relevant in the total cohort is reflective of the significant class imbalance between the positively and negatively predicted cohorts. LSTM, long short term memory; ECMO, extracorporeal membrane oxygenation; SHAP, Shapley Additive exPlanation.

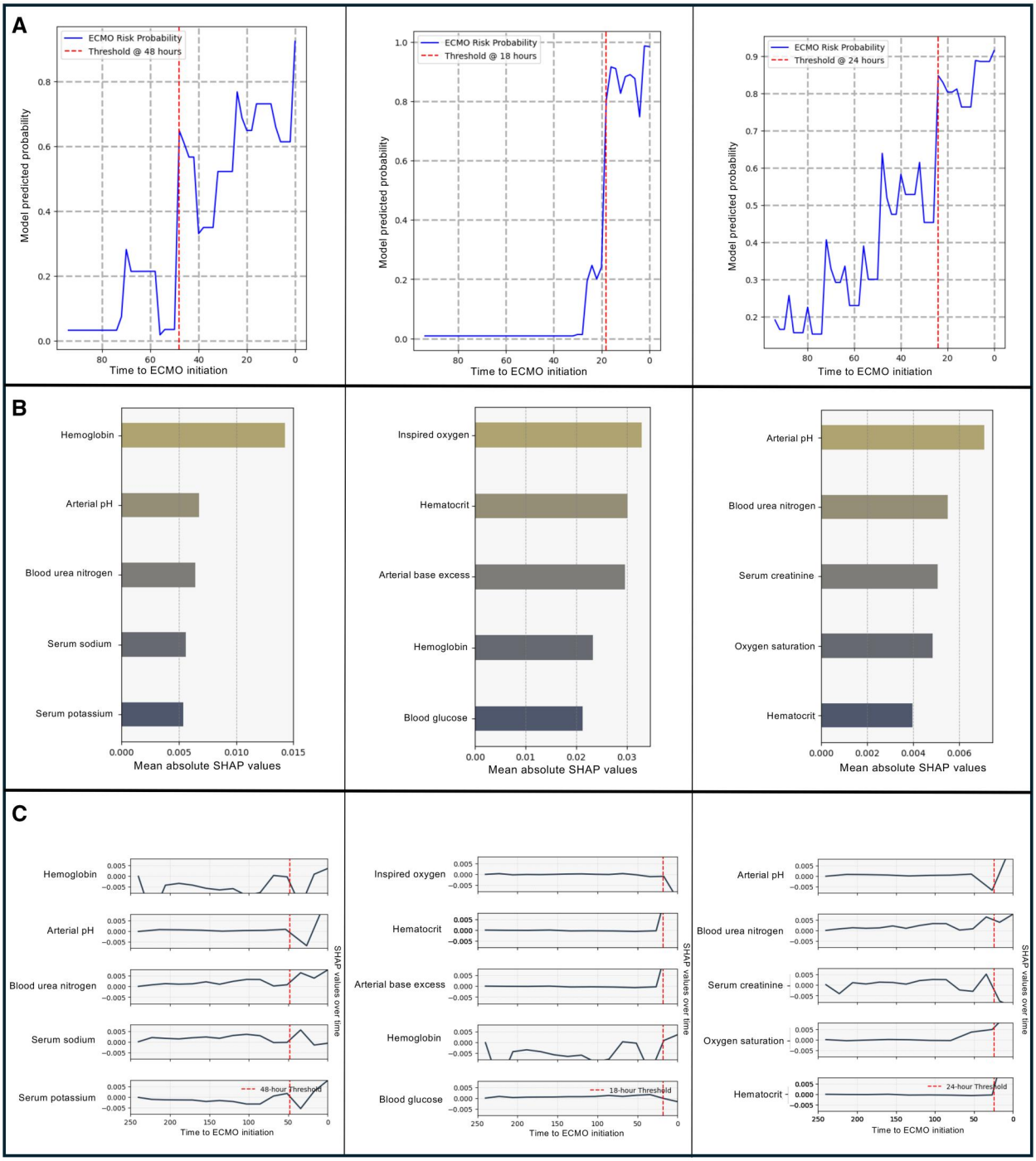

Figure 5. Change in PreEMPT-ECMO model performance and feature contribution over time by SHAP analysis. (A) Change in model performance per patient. Examples of 3 patients identified to be at high risk of receiving ECMO support. The model's predictive performance had an up-surge at varying timepoints in all 3 patients varying between, 48, 18, and 24 hours prior to actual ECMO use (time zero) as delineated by the red dotted line. (B) SHAP analysis of feature contribution per patient arranged by order of importance in the 3 different patients. The features identified to be most relevant in the model performance at the point of surge in model prediction varied by patient. While all features identified in all 3 patient examples have clinical relevance to ECMO decision making, the order of importance varied by patient highlighting the individualized predictive approach of the PreEMPT-ECMO model. (C) Change in feature contribution to model performance by SHAP analysis over time in the 3 different patients. The contribution of each feature to the model performance changed over time in each patient. The cumulative contribution of the top identified features led to a surge in model performance at varying time points prior to actual ECMO utilization, further illustrating the individualized nature of the model predictions, more accurately reflecting the ECMO decision making in real life. ECMO, extracorporeal membrane oxygenation; SHAP, SHapley Additive exPlanation.

References

1. Fernando SM, Qureshi D, Tanuseputro P, et al. Mortality and costs following extracorporeal membrane oxygenation in critically ill adults: a population-based cohort study. *Intensive Care Med*. 2019;45:1580-1589. https://doi.org/10.1007/s00134-019-05766-z [published Online First: 2019/09/19].
2. Mishra V, Svennevig JL, Bugge JF, et al. Cost of extracorporeal membrane oxygenation: evidence from the rikshospitalet university hospital, Oslo, Norway. *Eur J Cardiothorac Surg*. 2010;37:339-342. https://doi.org/10.1016/j.ejcts.2009.06.059 [published Online First: 2009/08/25].
3. Liao MT, Lin MH, Tsai HE, et al. Risk stratification and cost-effectiveness analysis of adult patients receiving extracorporeal membrane oxygenation. *J Eval Clin Pract*. 2022;28:615-623. https://doi.org/10.1111/jep.13681 [published Online First: 20220401].
4. Falcoz PE, Monnier A, Puyraveau M, et al. Extracorporeal membrane oxygenation for critically ill patients with COVID-19-related acute respiratory distress syndrome: worth the effort? *Am J Respir Crit Care Med*. 2020;202:460-463. https://doi.org/10.1164/rccm.202004-1370LE [published Online First: 2020/06/17].
5. Axiaq A, Haiduc AA, Alom S, Melamed N, Harky A. Extracorporeal membrane oxygenation in COVID-19: supplementary considerations. *J Card Surg*. 2020;35:3673-3674. https://doi.org/10.1111/jocs.15026 [published Online First: 2020/09/18].
6. Haiduc AA, Alom S, Melamed N, Harky A. Role of extracorporeal membrane oxygenation in COVID-19: a systematic review. *J Card Surg*. 2020;35:2679-2687. https://doi.org/10.1111/jocs.14879 [published Online First: 2020/07/28].
7. Supady A, Badulak J, Evans L, Curtis JR, Brodie D. Should we ration extracorporeal membrane oxygenation during the COVID-19 pandemic? *Lancet Respir Med*. 2021;9:326-328. https://doi.org/10.1016/S2213-2600(21)00131-4 [published Online First: 2021/04/03].
8. Zhang Y, Ji B, Zhou Z. ECMO support for COVID-19: a balancing act. *Lancet*. 2021;397:94-95. https://doi.org/10.1016/S0140-6736(20)32515-0 [published Online First: 2021/01/11].
9. Agerstrand C, Dubois R, Takeda K, et al. Extracorporeal membrane oxygenation for coronavirus disease 2019: crisis standards of care. *Asaio J*. 2021;67:245-249. https://doi.org/10.1097/MAT.0000000000001376 [published Online First: 2021/02/26].
10. Ramanathan K, Antognini D, Combes A, et al. Planning and provision of ECMO services for severe ARDS during the COVID-19 pandemic and other outbreaks of emerging infectious diseases. *Lancet Respir Med*. 2020;8:518-526. https://doi.org/10.1016/S2213-2600(20)30121-1 [published Online First: 2020/03/24].
11. Shekar K, Slutsky AS, Brodie D. ECMO for severe ARDS associated with COVID-19: now we know we can, but should we? *Lancet Respir Med*. 2020;8:1066-1068. https://doi.org/10.1016/S2213-2600(20)30357-X [published Online First: 2020/08/18].
12. Xue B, Shah N, Yang H, et al. Multi-horizon predictive models for guiding extracorporeal resource allocation in critically ill COVID-19 patients. *J Am Med Inform Assoc*. 2023;30:656-667. https://doi.org/10.1093/jamia/ocac256
13. Liu LJ, Ortiz-Soriano V, Neyra JA, Chen J. KIT-LSTM: knowledge-guided time-aware LSTM for continuous clinical risk prediction. *Proc (IEEE Int Conf Bioinform Biomed)*. 2022;2022:1086-1091. https://doi.org/10.1109/bibm55620.2022.9994931
14. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. *Nat Med*. 2022;28:1773-1784. https://doi.org/10.1038/s41591-022-01981-2
15. Landi I, Glicksberg BS, Lee H-C, et al. Deep representation learning of electronic health records to unlock patient stratification at scale. *NPJ Digit Med*. 2020;3:96. https://doi.org/10.1038/s41746-020-0301-z
16. Haendel MA, Chute CG, Bennett TD, N3C Consortium, et al. The national COVID cohort collaborative (N3C): rationale, design, infrastructure, and deployment. *J Am Med Inform Assoc*. 2021;28:427-443. https://doi.org/10.1093/jamia/ocaa196
17. Bennett TD, Moffitt RA, Hajagos JG, et al.; National COVID Cohort Collaborative (N3C) Consortium. Clinical characterization and prediction of clinical severity of SARS-CoV-2 infection among US adults using data from the US national COVID cohort collaborative. *JAMA Netw Open*. 2021;4:e2116901. https://doi.org/10.1001/jamanetworkopen.2021.16901 [published Online First: 20210701].
18. Stephens AF, Seman M, Diehl A, et al.; Extracorporeal Life Support Organization Member Centres. ECMO PAL: using deep neural networks for survival prediction in venoarterial extracorporeal membrane oxygenation. *Intensive Care Med*. 2023;49:1090-1099. https://doi.org/10.1007/s00134-023-07157-x [published Online First: 20230807].
19. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. https://doi.org/10.1136/bmj-2023-078378
20. Li D, Xue B, King C, et al. Self-explaining hierarchical model for intraoperative time series. In: *2022 IEEE International Conference on Data Mining (ICDM)*. IEEE; 2022.
21. Gao L, Li X, Song J, Shen HT. Hierarchical LSTMs with adaptive attention for visual captioning. *IEEE Trans Pattern Anal Mach Intell*. 2020;42:1112-1131. https://doi.org/10.1109/TPAMI.2019.2894139
22. Chung J, Ahn S, Bengio Y. Hierarchical multiscale recurrent neural networks. 2016. Accessed December 31, 2024. https://doi.org/10.48550/arXiv.1609.01704
23. Carvalho DV, Pereira EM, Cardoso JS. Machine learning interpretability: a survey on methods and metrics. *Electronics*. 2019;8:832.
24. Lundberg SA. Unified approach to interpreting model predictions. arXiv, arXiv:1705.07874, 2017, preprint. https://doi.org/10.48550/arXiv.1609.01704
25. Anai M, Akaike K, Iwagoe H, et al. Decrease in hemoglobin level predicts increased risk for severe respiratory failure in COVID-19 patients with pneumonia. *Respir Investig*. 2021;59:187-193. https://doi.org/10.1016/j.resinv.2020.10.009 [published Online First: 20201123].
26. Loughlin PC, Sebat F, Kellett JG. Respiratory rate: the forgotten vital sign - make it count! *Jt Comm J Qual Patient Saf*. 2018;44:494-499. https://doi.org/10.1016/j.jcjq.2018.04.014
27. Bulloch MN, Cardinale-King M, Cogle S, et al. Correction of electrolyte abnormalities in critically ill patients. *Intensive Care Res*. 2024;4:19-37. https://doi.org/10.1007/s44231-023-00054-3
28. Hayden SJ, Albert TJ, Watkins TR, Swenson ER. Anemia in critical illness: insights into etiology, consequences, and management. *Am J Respir Crit Care Med*. 2012;185:1049-1057. https://doi.org/10.1164/rccm.201110-1915CI [published Online First: 20120126].
29. Combes A, Brodie D, Bartlett R, International ECMO Network (ECMONet), et al. Position paper for the organization of extracorporeal membrane oxygenation programs for acute respiratory failure in adult patients. *Am J Respir Crit Care Med*. 2014;190:488-496. https://doi.org/10.1164/rccm.201404-0630CP
30. Peek GJ, Mugford M, Tiruvoipati R, CESAR Trial Collaboration, et al. Efficacy and economic assessment of conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure (CESAR): a multicentre randomised controlled trial. *Lancet*. 2009;374:1351-1363. https://doi.org/10.1016/s0140-6736(09)61069-2 [published Online First: 20090915].
31. Brodie D, Bacchetta M. Extracorporeal membrane oxygenation for ARDS in adults. *N Engl J Med*. 2011;365:1905-1914. https://doi.org/10.1056/NEJMct1103720
32. Savarimuthu S, BinSaeid J, Harky A. The role of ECMO in COVID-19: can it provide rescue therapy in those who are

critically ill? *J Card Surg*. 2020;35:1298-1301. https://doi.org/10.1111/jocs.14635 [published Online First: 2020/05/23].
33. Shekar K, Badulak J, Peek G, et al.; ELSO Guideline Working Group. Extracorporeal life support organization coronavirus disease 2019 interim guidelines: a consensus document from an international group of interdisciplinary extracorporeal membrane oxygenation providers. *Asaio J*. 2020;66:707-721. https://doi.org/10.1097/MAT.0000000000001193 [published Online First: 2020/07/01].
34. Group TCDHS. Does extracorporeal membrane oxygenation improve survival in neonates with congenital diaphragmatic hernia? *J Pediatr Surg*. 1999;34:720-725. https://doi.org/10.1016/S0022-3468(99)90363-9
35. Schlapbach LJ, Chiletti R, Straney L, et al.; Australian & New Zealand Intensive Care Society (ANZICS) Centre for Outcomes & Resource Evaluation (CORE) and the Australian & New Zealand Intensive Care Society (ANZICS) Paediatric Study Group. Defining benefit threshold for extracorporeal membrane oxygenation in children with sepsis-a binational multicenter cohort study. *Crit Care*. 2019;23:429. https://doi.org/10.1186/s13054-019-2685-1 [published Online First: 20191230].

JAMIA Open, 2025, 8, 1–11
https://doi.org/10.1093/jamiaopen/ooae158
Research and Applications