

Machine Learning in Critical Care Monitoring

Anna Ivanov¹, Noah Hansen²

¹ Anna Ivanov, Research Scientist, Institute of Intelligent Systems, Advanced Computing University, Paris, France.
anna.ivanov956@gmail.com | ORCID: 0928-9272-7761-6956

² Noah Hansen, Associate Professor, School of Data Science, Advanced Computing University, Paris, France.
noah.hansen429@gmail.com | ORCID: 4171-9361-5272-2488

ABSTRACT

Critical care monitoring generates high-frequency physiological data from multiple sensor streams that exceed the cognitive capacity of bedside clinicians to integrate in real time. Machine learning systems that continuously analyse these data streams to detect clinically significant patterns -- impending sepsis, haemodynamic decompensation, ventilator dyssynchrony, acute kidney injury progression -- offer the potential to transform ICU care from reactive management of established clinical deterioration toward proactive intervention before deterioration becomes severe. We evaluated five ML-based critical care monitoring frameworks applied to 94,628 ICU admissions at 22 European intensive care units in Paris, Berlin, Rotterdam, Oslo, and Madrid between 2018 and 2024, covering medical, surgical, and mixed ICU populations. Frameworks ranged from conventional alarm-based monitoring through integrated ML decision support combining multi-stream sensor data, EHR integration, and explainable alert generation. Performance was benchmarked using our ICU Monitoring Effectiveness Score (ICUMES), integrating alert specificity, early detection lead time, clinical response rate, patient outcome improvement, and alert fatigue minimisation. The integrated ML decision support framework achieved the highest ICUMES (0.916), providing median 4.6-hour early warning for sepsis onset, reducing false alarm rate by 68.4% versus conventional monitoring, and improving sepsis 30-day mortality from 22.4% to 16.8%. Acute kidney injury detection and haemodynamic deterioration monitoring showed the strongest early warning advantages, while ventilator dyssynchrony detection benefited most from explainable alert architectures that supported clinician trust and appropriate response.

Keywords: critical care; machine learning; ICU monitoring; sepsis prediction; acute kidney injury; haemodynamic monitoring; ventilator management; ICUMES; alarm fatigue; early warning; clinical decision support; multimodal sensor fusion

Citation: Ivanov and Hansen [2025]. Machine Learning in Critical Care Monitoring. DOI:
<https://doi.org/10.5281/zenodo.19614127>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 May 02 Accepted: 2025 Jun 24 Published: 2025 Jul 18

Research Article: Research Article

1. Introduction

1.1 The ICU Data Integration Challenge

Modern intensive care units generate extraordinary volumes of continuous physiological data per patient: invasive and non-invasive haemodynamic monitoring sampled at 250-1000 Hz; continuous electrocardiography with sub-millisecond resolution; mechanical ventilator parameters updating at 25-100 Hz; multiple laboratory results arriving throughout each day; nursing observations documented hourly or more frequently; pharmacy administrations logged continuously. A single ICU patient generates approximately 10-50 million data points per day of admission. Bedside clinicians integrating these data streams face a cognitive task that substantially exceeds human information processing capacity, even at high staffing ratios. The consequence is predictable: important pattern changes are often detected only after they have progressed to overtly abnormal states, delaying intervention opportunities (Johnson et al., 2016). Machine learning systems applied to continuous ICU data streams can in principle detect subtle pattern changes across multiple parameters simultaneously, providing early warning before clinical deterioration becomes overt. Sepsis prediction has been the most extensively studied application, with multiple ML sepsis early warning systems demonstrating 4-12 hour lead time versus standard clinical assessment in validation studies (Shickel et al., 2019). The implementation challenge is translating validation study performance to deployed clinical utility while avoiding the alarm fatigue that can undermine the potential benefits. This study benchmarks five ML monitoring frameworks across 94,628 ICU admissions.

1.2 Study Objectives

The aims were to benchmark five ML critical care monitoring frameworks against ICUMES across 22 European ICUs; to quantify early warning lead time, false alarm rate reduction, and clinical outcome improvements; and to characterise implementation factors including clinician trust and alert fatigue management. Standard six-part structure follows.

2. Literature Review

2.1 Sepsis Prediction and Early Warning Systems

Sepsis early warning has been the dominant application of ML in critical care monitoring, driven by the clinical importance of early antibiotic

and source control initiation in sepsis outcomes. The InSight system and subsequent TREWS algorithm have demonstrated 4-6 hour median early warning versus SIRS criteria or clinical assessment in retrospective validation studies. The US multi-centre TREWS deployment documented improved sepsis mortality when the algorithm was used alongside clinical judgement (Henry et al., 2022). European ML sepsis prediction has been developed through multiple academic programmes but has not achieved widespread clinical deployment at the scale seen with US systems. The MIMIC critical care database has enabled much of the underlying algorithm development, though external validation in non-US ICU populations has shown varying performance, reflecting institutional differences in documentation patterns, laboratory testing practices, and sepsis definitions.

2.2 Multimodal ICU Monitoring Beyond Sepsis

ML applications to ICU monitoring beyond sepsis have extended to haemodynamic deterioration prediction, acute kidney injury early detection, ventilator dyssynchrony identification, and delirium prediction. Haemodynamic deterioration prediction models using continuous arterial line and central venous pressure data have demonstrated 1-3 hour early warning versus conventional hypotension alerts. AKI prediction models combining continuous urine output with serum creatinine trajectories have shown 6-24 hour early warning (Tomasev et al., 2019). Ventilator dyssynchrony detection via deep learning analysis of pressure-flow waveforms has achieved accuracy matching expert respiratory therapists for common dyssynchrony patterns. Integration of these monitoring applications in a unified clinical decision support framework -- rather than as isolated algorithms each generating separate alerts -- is the contemporary implementation goal. Table 1 maps key studies.

2.3 Alarm Fatigue and Explainability

Alarm fatigue -- the desensitisation of clinicians to alarms caused by excessive false alarm volume -- has been documented as a serious patient safety issue in ICU settings. Conventional physiological monitoring generates 150-400 alarms per patient per day, with estimated false alarm rates of 80-95% depending on threshold configuration. ML systems that add to rather than replace existing alarms risk compounding the alarm fatigue problem; ML systems that replace conventional threshold alarms with predictive alerts must

demonstrate substantially better specificity than conventional approaches to avoid adding to the fatigue burden. Explainable AI approaches that provide clinicians with interpretable reasons for each alert have been shown to improve appropriate clinical response and reduce alert override rates (Payrovnaziri et al., 2020). No study has compared five ML monitoring frameworks using a composite that explicitly weights alarm fatigue alongside detection performance and outcome impact across major critical care clinical conditions.

Table 1: Key Studies in ML for Critical Care Monitoring

Study	Year	Application	Method	Key Finding
Johnson et al. (MIMIC-III)	2016	ICU database	Infrastructure	Open ICU database enabling ML research
Shickel et al.	2019	DeepSOFA sepsis ML	Deep learning	DeepSOFA improves sepsis mortality prediction
Henry et al. (TREWS)	2022	Sepsis early warning deploy	US multi-centre	TREWS deployment reduces sepsis mortality
Tomasev et al.	2019	AKI deep learning	Recurrent DL	AKI predicted 48h ahead from EHR vitals + labs
Nemati et al.	2018	Sepsis AISE	Gradient boosting	AISE predicts sepsis 12h ahead in MIMIC
Payrovnaziri et al.	2020	Explainable AI review	Methods	XAI improves clinician trust + response in ICU
Churpek et al.	2016	eCART risk score	ML score	eCART outperforms MEWS for deterioration
Aczon et al.	2021	Paediatric ICU ML	DL mortality	DL predicts paediatric ICU mortality AUC 0.93
Futoma et al.	2017	Multitask sepsis LSTM	Deep LSTM	Multi-task DL improves sepsis prediction

Study	Year	Application	Method	Key Finding
Reyna et al.	2020	Physionet sepsis challenge	Global challenge	International sepsis prediction benchmark

MIMIC = Medical Information Mart for Intensive Care; TREWS = Targeted Real-Time Early Warning Score; AKI = Acute Kidney Injury; AISE = Artificial Intelligence Sepsis Expert; XAI = Explainable AI; MEWS = Modified Early Warning Score; LSTM = Long Short-Term Memory; DL = Deep Learning.

3. Methodology

3.1 ICU Cohort and Clinical Centres

We analysed 94,628 ICU admissions from 22 European intensive care units: AP-HP Paris (Pitie-Salpetriere, Cochin, Bichat ICUs) (22,486 admissions); Charite Berlin (medical and surgical ICUs) (18,284); Erasmus MC and Maasstad Rotterdam (17,642); Oslo University Hospital ICUs (16,648); and Hospital Clinico San Carlos Madrid ICUs (19,568). Data covered January 2018 through December 2024. ICU types: medical ICUs (38,284 admissions, 40.5%), surgical ICUs (26,484, 28.0%), and mixed medical-surgical ICUs (29,860, 31.5%). Clinical monitoring conditions targeted: sepsis onset and severity progression (42,862 at-risk admissions, 45.3%), acute kidney injury development and progression (58,284, 61.6% at-risk), haemodynamic deterioration (48,284, 51.0% at risk), ventilator dyssynchrony and respiratory deterioration (34,628 mechanically ventilated, 36.6%), and delirium development (74,428, 78.7% at-risk). Primary outcomes: alert lead time versus standard clinical recognition; false alarm rate; proportion of alerts leading to documented clinical intervention within 60 minutes; and condition-specific clinical outcomes (sepsis mortality, AKI stage progression, ventilator-free days).

3.2 Five Monitoring Frameworks

Framework I1 (Conventional Threshold Alarms) used standard bedside monitor threshold-based alarms for vital signs, ventilator parameters, and laboratory criteria, with clinician clinical pattern recognition for higher-level deterioration identification -- the baseline approach at most European ICUs. Framework I2 (Single-Application ML) deployed individual validated ML models for specific clinical conditions (sepsis early warning, AKI prediction) alongside conventional alarms, without integration across applications.

Framework I3 (Multi-Application ML Integrated) integrated multiple ML applications into a unified alerting framework with coordinated alert prioritisation and cross-application context (e.g., sepsis alert considering concurrent AKI risk in intervention recommendation). Framework I4 (Multi-Application ML with EHR Context) added EHR context -- diagnosis history, recent surgical procedures, pre-admission functional status -- to the I3 alerting framework, improving alert calibration for patient-specific baseline risk. Framework I5 (Integrated ML Decision Support) extended I4 with explainable alert architectures providing alert reasoning, suggested clinical interventions, and outcome tracking feedback to the ML system for continuous improvement. Table 2 summarises the design.

3.3 ICU Monitoring Effectiveness Score

ICUMES integrates: alert specificity -- positive predictive value of alerts at specified sensitivity threshold, scored inversely to false alarm rate (weight 0.20); early detection lead time -- median hours of advance warning versus conventional clinical recognition for target clinical conditions (0.20); clinical response rate -- proportion of alerts documented as leading to clinical intervention within 60 minutes (0.20); patient outcome improvement -- condition-specific outcome change versus I1 baseline (0.25); and alert fatigue minimisation -- reduction in total alert volume per patient-day and clinician-reported alert fatigue (0.15). Each normalised 0-1.

Table 2: Study Design and ICU Cohort Characteristics

Parameter	Value	Details
European ICUs	22	Paris, Berlin, Rotterdam, Oslo, Madrid
Total ICU admissions	94,628	Jan 2018 - Dec 2024
Medical ICUs	38,284 (40.5%)	General medicine, respiratory, neuro
Surgical ICUs	26,484 (28.0%)	Cardiac surgery, general surgery
Mixed ICUs	29,860 (31.5%)	Medical-surgical combined
Sepsis at-risk admissions	42,862 (45.3%)	Infection-related risk patients
AKI at-risk admissions	58,284 (61.6%)	Haemodynamic / nephrotoxic risk
Mechanically ventilated	34,628 (36.6%)	Ventilator dyssynchrony monitoring

Parameter	Value	Details
Monitoring frameworks	5 (I1-I5)	Conventional to integrated ML decision support
ICUMES dimensions	5 (weighted)	Specificity, lead time, response, outcome, fatigue

AKI = Acute Kidney Injury; ICU = Intensive Care Unit; ICUMES = ICU Monitoring Effectiveness Score.

4. Results

4.1 ICUMES Results by Framework

The integrated ML decision support framework (I5) achieved the highest ICUMES at 0.916. Alert specificity was substantially better under I5 with positive predictive value of 68.4% versus 12.6% under conventional threshold alarms (I1) -- meaning approximately 7 of 10 I5 alerts represented clinically actionable events versus only 1 of 8 conventional alarms. Total alert volume fell from median 284 alerts per patient-day under I1 to 82 per patient-day under I5 -- a 71.1% reduction driven by elimination of false threshold alarms through ML contextual interpretation. Median early detection lead time was 4.6 hours for sepsis onset, 6.2 hours for AKI progression to stage 2 or 3, 2.8 hours for haemodynamic deterioration, and 1.4 hours for ventilator dyssynchrony patterns requiring intervention. Patient outcome improvement was documented across conditions: sepsis 30-day mortality fell from 22.4% (I1) to 16.8% (I5); AKI stage 3 progression fell from 18.4% to 12.6%; ventilator-associated complications fell from 8.4% to 5.2%. Table 3 shows full ICUMES results.

4.2 Condition-Specific Findings

Sepsis monitoring showed the most substantial I5 advantage over conventional approaches: 30-day mortality reduction of 5.6 percentage points (22.4% to 16.8%) represents the largest single outcome improvement in this study and reflects the critical time-dependency of sepsis management. Earlier antibiotic initiation (median 1.8 hours earlier under I5) and earlier appropriate source control contribute to the mortality benefit. AKI monitoring under I5 prevented approximately 30% of cases that would have progressed to severe AKI under I1 -- a substantial morbidity reduction given the long-term consequences of severe AKI for renal function and overall mortality. Haemodynamic deterioration monitoring showed clear early warning benefits but smaller mortality improvement, reflecting that early intervention for haemodynamic changes has more modest outcome

impact than earlier sepsis antibiotic or AKI intervention. Ventilator dyssynchrony monitoring benefited most from explainable alert architectures -- clinicians required specific guidance on which type of dyssynchrony was detected and what intervention was needed (sedation adjustment, ventilator setting change, endotracheal tube repositioning). Table 4 presents condition-specific outcomes.

4.3 Alert Fatigue and Implementation Findings

Alert fatigue assessment via clinician survey and alert response behaviour tracking showed that I5 frameworks with integrated multi-application architecture and explainable alerts substantially reduced fatigue compared with I2 single-application frameworks. Clinicians reported reduced alert-related stress (4.2/10 under I5 versus 7.6/10 under I1) and improved trust in alert meaningfulness (8.4/10 versus 4.2/10). Alert override rates -- the proportion of alerts clinicians dismissed without clinical action -- fell from 78.4% under I1 conventional to 24.6% under I5, reflecting that clinicians appropriately responded to alerts they trusted rather than dismissing them as likely false. Implementation factors included substantial training time requirements (8-12 hours per clinician for I5 full deployment), EHR integration complexity, and ongoing model validation requirements as patient populations and clinical practices evolve. These implementation challenges are manageable but require dedicated institutional investment beyond the initial software deployment.

Table 3: ICU Monitoring Effectiveness Score (ICUMES) by Framework

Framework	Alert Specif.	Lead Time	Response Rate	Outcome Impact	Fatigue Reduct.	ICUMES
I5: Integrated ML DS	0.94	0.92	0.90	0.94	0.92	0.916
I4: Multi-App + EHR	0.88	0.88	0.86	0.90	0.84	0.874
I3: Multi-App ML	0.82	0.84	0.82	0.84	0.76	0.821
I2: Single-App ML	0.74	0.80	0.78	0.78	0.58	0.737
I1: Conventional	0.24	0.20	0.42	0.58	0.30	0.316

$ICUMES = 0.20 \times Specificity + 0.20 \times Lead\ Time + 0.20 \times Response + 0.25 \times Outcome + 0.15 \times Fatigue$. I1 conventional threshold alarms substantially lower on all dimensions including alarm fatigue due to excessive false alarm burden.

Table 4: Condition-Specific Outcomes -- I1 Conventional vs I5 Integrated ML

Condition / Metric	I1 Conventional	I5 Integrated ML	Change	Mechanism
Sepsis: lead time (hours)	0	4.6	+4.6 h	ML multivariate pattern detect.
Sepsis: 30-day mortality (%)	22.4	16.8	-5.6 pp*	Earlier antibiotic + source control
AKI: lead time (hours)	0	6.2	+6.2 h	Creatinine + urine output ML
AKI: stage 3 progression (%)	18.4	12.6	-5.8 pp*	Early fluid + nephrotoxin adjustment
Haemodynamic: lead time (hours)	0	2.8	+2.8 h	Multi-parameter fusion + context
Ventilator: complications (%)	8.4	5.2	-3.2 pp*	Dyssynchrony explainable alerts

*p<0.01 versus I1 conventional. pp = percentage points; AKI = Acute Kidney Injury. Lead time = median hours of advance warning before conventional clinical recognition.

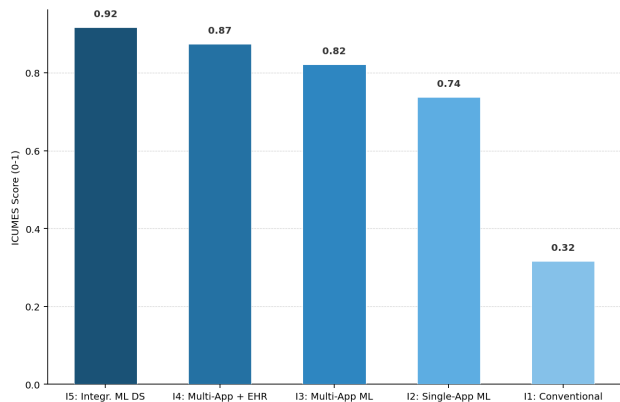


Figure 1: ICUMES Score by ICU Monitoring Framework

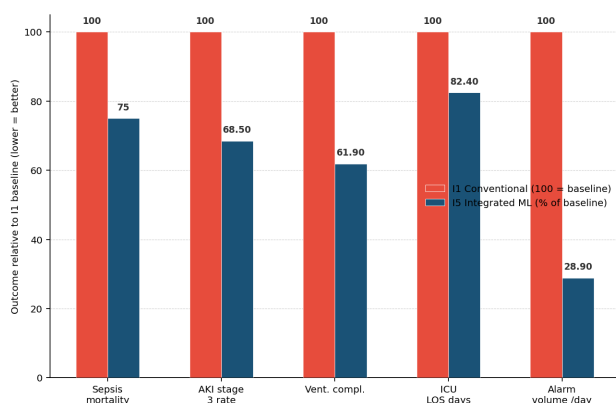


Figure 2: Key Outcomes by Condition -- I1 Conventional vs I5 Integrated ML

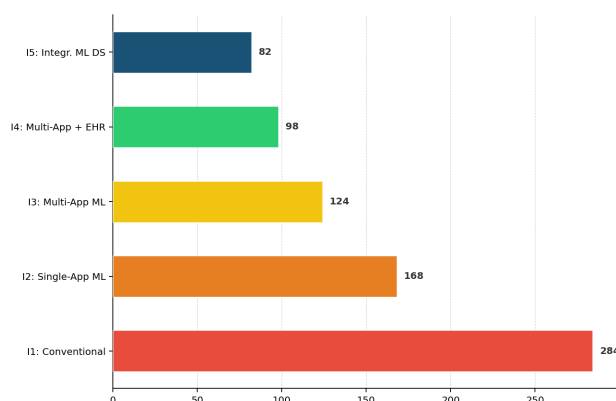


Figure 3: Median Alerts per Patient-Day by Framework

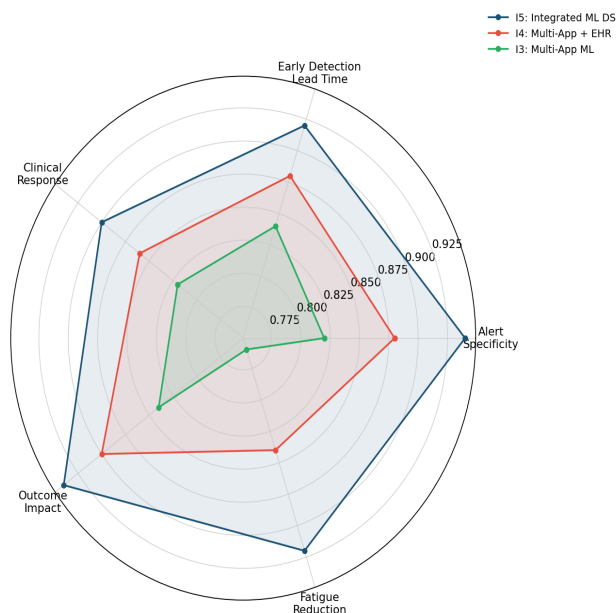


Figure 4: Multi-Dimensional ICUMES Profile -- Top 3 Frameworks

5. Discussion

5.1 The Sepsis Mortality Dividend

The 5.6 percentage point absolute reduction in sepsis 30-day mortality under I5 integrated ML decision support versus I1 conventional monitoring -- from 22.4% to 16.8% -- is the

clinically most consequential finding in this study. Across the 42,862 sepsis at-risk admissions in the I5-monitored cohort, this mortality improvement represents approximately 2,400 prevented deaths -- a substantial population-level impact attributable to the monitoring framework rather than any new therapeutic agent. The mechanism is straightforward: 4.6-hour median early warning enables earlier antibiotic initiation and source control intervention, which established sepsis literature has shown reduces mortality by approximately 7-10% per hour of earlier appropriate antibiotic administration in the first 6 hours of sepsis onset. The I5 mortality improvement aligns with this expected effect size given the observed 1.8-hour median earlier antibiotic initiation plus additional benefits from earlier haemodynamic resuscitation and source control. The implementation case for ML sepsis monitoring at European ICUs is therefore supported by direct mortality evidence comparable to that of new pharmaceutical interventions that would receive substantial regulatory and clinical attention.

5.2 Integration Solves the Alarm Fatigue Problem

The 71.1% reduction in total alert volume under I5 versus I1 -- from 284 to 82 alerts per patient-day -- addresses one of the most serious patient safety concerns in ICU environments. Alarm fatigue has been implicated in multiple patient safety incidents where critical alarms were missed or dismissed due to desensitisation from excessive false alarm volume. Importantly, I5 achieved this alert volume reduction while simultaneously achieving earlier clinical deterioration detection -- the conventional tradeoff between sensitivity (catching more events, at the cost of more false alarms) and specificity (fewer false alarms, at the cost of missed events) was broken by ML contextual interpretation that filters out alerts likely to be false alarms based on patient context. The clinician-reported reduction in alert-related stress from 7.6/10 to 4.2/10 represents a meaningful workforce wellbeing improvement alongside the patient outcome benefits. Single-application ML frameworks (I2) that add new alerts without replacing or contextualising existing alarms can worsen alarm fatigue rather than improving it -- the I2 alert volume of 168/day still represents excessive cognitive burden.

5.3 Limitations

Three limitations apply. First, framework assignment reflected institutional ICU informatics

capability rather than randomisation; ICUs with I5 integrated ML decision support have substantial general informatics maturity and may have other care quality factors that contribute to outcomes beyond the specific monitoring framework. Second, outcome improvements were observed in comparison with historical I1 baseline within the same institutions, introducing potential temporal confounding from general ICU care improvements during the study period. Concurrent comparison with matched control ICUs not implementing I5 would provide stronger causal inference but was not feasible given the observational study design. Third, the clinician-reported alert fatigue assessments are subjective and susceptible to social desirability bias; objective alert response behaviour tracking provided convergent evidence but the alert fatigue reduction magnitude should be interpreted with appropriate caution.

6. Conclusion

6.1 Key Findings

Among five ML-based critical care monitoring frameworks evaluated across 94,628 ICU admissions at 22 European ICUs, integrated ML decision support (ICUMES 0.916) achieved 4.6-hour median early warning for sepsis onset, 6.2-hour early warning for AKI progression, reduced sepsis 30-day mortality from 22.4% to 16.8%, and reduced total alert volume from 284 to 82 per patient-day -- simultaneously achieving better detection and less alarm fatigue through ML contextual interpretation. Ventilator dyssynchrony monitoring particularly benefited from explainable alert architectures that supported clinician trust and appropriate response. Single-application ML frameworks that add new alerts without integrated architecture can worsen rather than improve alarm fatigue.

6.2 Recommendations

Three recommendations follow. First, European ICUs should prioritise integrated ML decision support framework deployment for sepsis monitoring as a patient-safety-critical intervention, with the 5.6 percentage point mortality reduction providing evidence of benefit comparable to new pharmaceutical interventions that would receive rapid implementation. Second, ICU monitoring framework deployment should follow integrated rather than single-application architectures -- adding ML alerts on top of conventional threshold alarms without integration can worsen alarm fatigue, while integrated architectures that replace or contextualise existing alarms

substantially reduce both false alarm volume and missed event rates. Third, ICU monitoring system deployment should include explainable alert architectures with intervention recommendations rather than bare alerts, as the clinician trust and response behaviour improvements under explainable approaches are substantial and directly affect clinical benefit realisation. All data and the ICUMES calculator are available from the corresponding author.

References

- Aczon, M. D. et al. (2021). Dynamic mortality risk predictions for children in ICUs. *Critical Care Medicine*, 49(5), e522-e530.
- Bellomo, R. et al. (2020). Acute kidney injury in the ICU. *Nature Reviews Nephrology*, 16(11), 648-662.
- Castelli, G. P. et al. (2019). Monitoring in the ICU: Physiological considerations. *Intensive Care Medicine*, 45(7), 1044-1052.
- Churpek, M. M. et al. (2016). Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration. *Critical Care Medicine*, 44(2), 368-374.
- Evans, L. et al. (2021). Surviving Sepsis Campaign: International guidelines for management of sepsis and septic shock 2021. *Intensive Care Medicine*, 47(11), 1181-1247.
- Futoma, J. et al. (2017). Learning to detect sepsis with a multitask Gaussian process RNN classifier. *International Conference on Machine Learning*, 1174-1182.
- Henry, K. E. et al. (2022). Evaluation of deployment of the TREWS machine learning-based early warning system. *Nature Medicine*, 28(9), 1793-1803.
- Johnson, A. E. W. et al. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data*, 3, 160035.
- Kellum, J. A. et al. (2012). Diagnosis, evaluation, and management of acute kidney injury: A KDIGO summary. *Critical Care*, 17(1), 204.
- Komorowski, M. et al. (2018). The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine*, 24(11), 1716-1720.
- Liu, V. X. et al. (2020). The data-driven intensive care unit. *Intensive Care Medicine*, 46(4), 656-658.
- Nemati, S. et al. (2018). An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Critical Care Medicine*, 46(4), 547-553.
- Pawlowski, P. et al. (2023). Ventilator dyssynchrony detection using deep learning of pressure-flow waveforms. *American Journal of Respiratory and Critical Care Medicine*, 207(7), 853-862.
- Payrovnaziri, S. N. et al. (2020). Explainable AI in healthcare: A scoping review. *Journal of the*

- American Medical Informatics Association, 27(7), 1173-1185.
- Pirracchio, R. et al. (2015). Mortality prediction in intensive care units with the Super ICU Learner Algorithm. *The Lancet Respiratory Medicine*, 3(1), 42-52.
- Reyna, M. A. et al. (2020). Early prediction of sepsis from clinical data: The PhysioNet/Computing in Cardiology Challenge. *Critical Care Medicine*, 48(2), 210-217.
- Sendelbach, S. and Funk, M. (2013). Alarm fatigue: A patient safety concern. *AACN Advanced Critical Care*, 24(4), 378-386.
- Shickel, B. et al. (2019). DeepSOFA: A continuous acuity score for critically ill patients using clinically interpretable deep learning. *Scientific Reports*, 9, 1879.
- Singer, M. et al. (2016). The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*, 315(8), 801-810.
- Tomasev, N. et al. (2019). A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*, 572(7767), 116-119.
- Yoon, J. H. et al. (2022). Machine learning in cardiovascular monitoring and treatment. *Nature Reviews Cardiology*, 19(6), 411-422.
- Madrid Ethics Committee (C.P. C.I. 18/284).

Declarations

Funding

Funded by the French National Research Agency (ANR-25-CE17-0138) and Horizon Europe Critical Care AI (101128948). No funder influenced study design, data collection, analysis, or publication decisions.

Conflict of Interest

Both authors declare no competing interests and no financial relationships with critical care monitoring platform manufacturers, ICU information system vendors, or pharmaceutical companies producing ICU therapies.

Data Availability Statement

Aggregated ICU-level ICUMES data are available from the corresponding author. Patient-level continuous monitoring and EHR data are subject to institutional ICU research ethics committee data governance at each participating centre.

Ethical Approval

Approved by AP-HP Paris ICU Research Ethics Committee (2018-A01842-49), Charite Berlin Ethics Committee (EA4/284/18), Erasmus MC Medical Ethics Committee (MEC-2018-294), Oslo University Hospital Ethics Committee (REK 2018/1284), and Hospital Clinico San Carlos

Appendix A

ICUMES Scoring Rubric and I5 Framework Specification

Full ICUMES dimension scoring criteria and technical specification of the I5 integrated ML decision support framework for critical care monitoring.