ISSN: 1607-8322, e-ISSN: 2220-5799 Anaesthesia, Pain & Intensive Care
Vol 29(6); September 2025 DOI: 10.35975/apic.v29i6.2908

CORRESPONDENCE INTENSIVE CARE

Recognizing the boundaries of innovation in the Intensive Care Unit

Faraz Mansoor*

Faraz Mansoor, Consultant Anesthesiologist, Shaukat Khanum Memorial Cancer Hospital & Research Center, Phase 5, Hayatabad, Peshawar, Pakistan; Email: mansoor_faraz@hotmail.com; {ORCID:0000-0002-7477-3263}

I read with great interest the recent editorial highlighting the role of artificial intelligence (AI) in the intensive care unit (ICU).[1] While AI has generated considerable enthusiasm in critical care research, several key publications underscore its current shortcomings and the need for a more cautious approach to implementation.

The experience with the 'Epic Sepsis Model' (ESM) illustrates how insufficiently validated AI tools can underperform in real-world ICU settings, producing low sensitivity, poor calibration, and frequent false alarms that undermine clinical utility.[2] Tools like the ESM have faced criticism for the same reasons, leading to alert fatigue. Surprisingly, its external validation at the University of Michigan revealed only a 33% sensitivity and an AUC of 0.63—significantly lower than manufacturer-reported values (AUC 0.76–0.83).[3] Furthermore, a follow-up analysis in NEJM AI reported that the ESM's performance dropped further when evaluated for predictions made strictly before clinical recognition—AUROC fell from 0.62 to 0.47—suggesting that the model may rely on indicators of clinician suspicion rather than truly early detection.[4] On top of this, outside the academic centers, for example, implementations at county hospital emergency department, produced poor metrics: only 14.7% sensitivity, 95.3% specificity, and a median lead time of zero minutes—indicating alerts often arrived too late or not at all.[5] Beyond the shortcomings of individual tools, like the ESM, these challenges reflect a deeper limitation of AI in critical care—its heavy dependence on the quality and scope of underlying datasets, which often lack the generalizability needed for diverse ICU environments.

The AI models are data-dependent and lack generalizability. For example, the MIMIC-III ('Medical Information Mart for Intensive Care-III') database, which is a freely available, single-center critical care dataset with comprehensive patient data (vitals, labs, medications, notes) from over forty thousand ICU admissions, has been instrumental in developing and validating numerous prediction models. However, because MIMIC-III originates from a single tertiary center, models trained on it often lack external validity, limiting generalizability.[6] In other words, these models might not be suitable for a tertiary center ICU in Pakistan.

Several notable reviews underscore the current limitations of AI in critical care. Luo et al. (2022) emphasized that despite advances in sepsis management and mechanical ventilation, AI applications remain constrained by significant risks of bias, limited generalizability, and unresolved ethical concerns, illustrating that AI is "far from perfect" in ICU settings.[7] Likewise, Lovejoy et al. (2019) highlighted the dearth of positive multicenter prospective randomized controlled trials and the challenges posed by heterogeneous, incomplete data—which compromise the reliability of AI tools in diverse ICU environments.[8] In the domain of mechanical ventilation, Misseri et al. (2024) described AI's role as still nascent, urging a cautious and measured approach toward its clinical integration.[9] Furthermore, Godbole et al. pointed out a lack of robust, real-world studies applying AI to infection prevention and control in ICU practice—underscoring the persistent gap between algorithmic potential and clinical implementation.[10]

Taken together, these studies converge on a central message: although AI holds significant promise, its present clinical value in the ICU is limited by methodological flaws, poor external validity, workflow integration barriers, and ethical concerns. Future progress will require carefully designed, prospective multicentre evaluations and a stronger emphasis on interpretability, transparency, and clinician–AI collaboration. Until such evidence is available, widespread adoption of AI in intensive care should remain measured and cautious.

The future of AI in intensive care lies in careful, evidence-based integration rather than premature adoption. Robust multicenter validation and explainable models are essential to ensure reliability and foster

clinician trust, while user-centered design must minimize workflow disruption and alert fatigue. Ethical and legal frameworks addressing accountability, consent, and privacy are equally critical to safeguard the patients and the healthcare providers.

Finally, AI should be seen not as a replacement, but as an adjunct—an "augmented intelligence" that strengthens, rather than supplants, the art and science of critical care.

Conflict of interest

The author claims no conflict of interests.

Authors contribution

FM is the sole author of this paper.

REFERENCES

1. Naseer A, Kiran Q. Artificial intelligence – new horizons in intensive care. Anaesth Pain Intensive Care. 2025;29(5):254-6. Full Text
2. Lyons P, Ramsey B, Simkins M, Maddox T. How useful is the Epic Sepsis Prediction Model for predicting sepsis? Am J Respir Crit Care Med. 2021;203(1):A1580. DOI: 10.1164/ajrccm-conference.2021.203.1_MeetingAbstracts.A1580
3. Wong A, Ötles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. JAMA Intern Med. 2021;181(8):1065-1070. PubMed
4. Kamran F, Tjandra D, Heiler A, Virzi JM, Singh K, King JE, et al. Evaluation of sepsis prediction models before onset of treatment. NEJM AI. 2024;1(3). DOI:10.1056/AIoa2300032
5. Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, et al. Prospective, multisite study of patient outcomes after implementation of the Epic sepsis prediction model. Crit Care Med. 2022;50(9):1322-31. DOI: 10.1038/s41591-022-01894-0
6. Johnson AEW, Pollard TJ, Lehman LWH, et al. MIMIC-III, a freely accessible critical care database. Crit Care Med. 2016;44(5):774-81. PubMed DOI: 10.1038/sdata.2016.35
7. Luo MH, Huang DL, Luo JC, Su Y, Li JK, Tu GW, et al. Data science in the intensive care unit. World J Crit Care Med. 2022;11(5):311-6. PubMed DOI: 10.5492/wjccm.v11.i5.311
8. Lovejoy CA, Buch V, Maruthappu M. Artificial intelligence in the intensive care unit. Crit Care. 2019;23:7. PubMed DOI: 10.1186/s13054-018-2301-9
9. Misseri G, Piattoli M, Cuttone G, Gregoretti C, Bignami EG. Artificial intelligence for mechanical ventilation: a transformative shift in critical care. Ther Adv Pulm Crit Care Med. 2024;19:29768675241298918. PubMed DOI: 10.1177/29768675241298918
10. Godbole AA, Paras, Mehra M, Banerjee S, Roy P, Deb N, et al. Enhancing infection control in ICUs through AI: a literature review. Health Sci Rep. 2025;8(1):e70288. PubMed DOI: 10.1002/hsr2.70288

