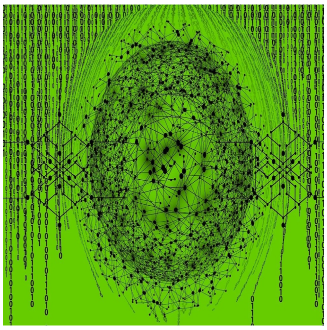
ICU MANAGEMENT & PRACTICE

Deep Learning Models for ICU Readmission Prediction



The decision to discharge patients from the ICU is complex, as premature discharge increases the risk of readmission and mortality, while delayed discharge raises healthcare costs. ICU readmission is associated with over twice the mortality risk, longer hospital stays, greater morbidity, and higher resource use. Several predictive models, such as SWIFT, APACHE II and III, and SAPS II, have been developed to estimate readmission risk, but their clinical utility is limited due to poor real-world performance, methodological biases, and reliance on restrictive statistical assumptions that fail to capture complex, high-dimensional data.

Machine learning (ML) approaches, particularly deep learning (DL), can model nonlinear relationships and interactions among multiple variables, offering potential advantages for ICU readmission prediction. While DL has shown promise in medical imaging and hospital outcome prediction, its application and validation in ICU readmission remain underexplored. Existing reviews largely overlook DL-specific models.

A recent review examines the current landscape of DL-based ICU readmission prediction methods, evaluating their performance, identifying key challenges, and outlining future directions to inform clinical adoption in critical care settings. The review developed or validated DL models for predicting ICU readmission, using PubMed, Embase, Scopus, and Web of Science. The studies were analysed based on outcome and population definitions, DL architectures, reproducibility, generalisability, and explainability, and a meta-analysis was performed to estimate overall model performance.

The review included 24 studies comprising 49 DL models. Studies showed wide variability in ICU readmission definitions, timeframes, and DL architectures, with a high risk of bias and limited reproducibility or interpretability. Meta-analysis of 11 studies reported a pooled AUROC of 0.78 (95% CI = 0.72–0.84) with extreme heterogeneity ($I^2 = 99.9\%$). Models focused on disease-specific ICU populations performed significantly better, achieving a mean AUROC of 0.92 (95% CI = 0.89–0.95) and much lower heterogeneity ($I^2 = 17.1\%$).

Accurate prediction of ICU readmissions is essential to prevent adverse outcomes and reduce healthcare costs. DL models show promise in improving prediction accuracy, achieving a median AUROC increase of 11% over traditional models and demonstrating potentially better generalisability. However, several challenges limit their clinical applicability.

Technical reproducibility was poor; only three studies shared code despite using publicly available datasets. High computational demands further restrict accessibility and real-time deployment. Conceptual reproducibility was also limited, as most models were trained on single-centre U.S. datasets, showing bias and inconsistent performance across settings. Only two studies used non-U.S. data, and few conducted external validation, which revealed reduced performance. The lack of multi-institutional datasets and overfitting to training data further hinder generalisability. Statistical reproducibility was also weak, with many studies failing to report variance metrics.

Substantial variability existed across studies in datasets, outcome definitions, timeframes, and DL architectures, contributing to high heterogeneity and risk of bias. Disease- or ICU-specific models performed better due to more homogeneous patient groups, but comparisons between DL methods were scarce.

Overall, while DL models show encouraging predictive performance for ICU readmission, issues with reproducibility, generalisability, heterogeneity, and explainability must be addressed before clinical integration.

Source: Critical Care

Image Credit: Gerd Altmann from Pixabay

Published on : Tue, 21 Oct 2025