

Anesthesiology. 2022 Dec 20. doi: 10.1097/ALN.0000000000004478. Online ahead of print.

Predicting Intensive Care Delirium with Machine Learning: Model Development and External Validation

Kirby D Gong ¹, Ryan Lu ², Teya S Bergamaschi ³, Akaash Sanyal ⁴, Joanna Guo ⁵, Han B Kim ⁶, Hieu T Nguyen ⁷, Joseph L Greenstein ⁸, Raimond L Winslow ⁹, Robert D Stevens ¹⁰

Affiliations

PMID: 36538354 DOI: [10.1097/ALN.0000000000004478](https://doi.org/10.1097/ALN.0000000000004478)

Abstract

Background: Delirium poses significant risks to patients, but countermeasures can be taken to mitigate negative outcomes. Accurately forecasting delirium in ICU patients could guide proactive intervention. Our primary objective was to predict ICU delirium by applying machine learning to clinical and physiological data routinely collected in electronic health records.

Methods: Two prediction models were trained and tested using a multi-center database (years of data collection 2014–15), and externally validated on two single-center databases (2001–2012 and 2008–2019). The primary outcome variable was delirium defined as a positive Confusion Assessment Method for the ICU screen, or an Intensive Care Delirium Screening Checklist ≥ 4 . The first model, named "24-hour model", used data from the 24 hours following ICU admission to predict delirium any time afterwards. The second model designated "dynamic model", predicted the onset of delirium up to 12 hours in advance. Model performance was compared to results using features from a widely-cited reference model.

Results: For the 24-hour model, delirium was identified in 2,536/18,305 (13.9%), 768/5,299 (14.5%), and 5,955/36,194 (11.9%) of patient stays respectively in the development sample and two validation samples. For the 12-hour lead time dynamic model, delirium was identified in 3,791/22,234 (17.0%), 994/6,166 (16.1%), and 5,955/28,440 (20.9%) patient stays, respectively. Mean AUC (95% CI) for the first 24-hour model was 0.785 (0.769, 0.801), significantly higher than the modified reference model with AUC of 0.730 (0.704, 0.757). The dynamic model had a mean AUC of 0.845 (0.831, 0.859) when predicting delirium 12 hours in advance. Calibration was similar in both models (mean Brier Score [95% CI] 0.102 [0.097, 0.108] and 0.111 [0.106, 0.116]). Model discrimination and calibration were maintained when tested on the validation datasets.

Conclusions: Machine learning models trained with clinical and physiological data predict ICU delirium and support dynamic time-sensitive forecasting.

Copyright © 2022, the American Society of Anesthesiologists. All Rights Reserved.