Predicting Intensive Care Unit Readmission with Machine Learning Using Electronic Health Record Data

Juan C. Rojas[1,2], Kyle A. Carey[1], Dana P. Edelson[1,2], Laura R. Venable[1], Michael D. Howell[1,2,3], and Matthew M. Churpek[1,2]

[1]Department of Medicine and [2]The Center for Healthcare Delivery Science and Innovation, University of Chicago, Chicago, Illinois; and [3]Google Research, Mountain View, California

Abstract

Rationale: Patients transferred from the intensive care unit to the wards who are later readmitted to the intensive care unit have increased length of stay, healthcare expenditure, and mortality compared with those who are never readmitted. Improving risk stratification for patients transferred to the wards could have important benefits for critically ill hospitalized patients.

Objectives: We aimed to use a machine-learning technique to derive and validate an intensive care unit readmission prediction model with variables available in the electronic health record in real time and compare it to previously published algorithms.

Methods: This observational cohort study was conducted at an academic hospital in the United States with approximately 600 inpatient beds. A total of 24,885 intensive care unit transfers to the wards were included, with 14,962 transfers (60%) in the training cohort and 9,923 transfers (40%) in the internal validation cohort. Patient characteristics, nursing assessments, *International Classification of Diseases, Ninth Revision* codes from prior admissions, medications, intensive care unit interventions, diagnostic tests, vital signs, and laboratory results were extracted from the electronic health record and used as predictor variables in a gradient-boosted machine model. Accuracy for predicting intensive care readmission was compared with the Stability and Workload Index for Transfer score and Modified Early Warning Score in the internal

validation cohort and also externally using the Medical Information Mart for Intensive Care database ($n = 42{,}303$ intensive care unit transfers).

Results: Eleven percent (2,834) of discharges to the wards were later readmitted to the intensive care unit. The machine-learning–derived model had significantly better performance (area under the receiver operating curve, 0.76) than either the Stability and Workload Index for Transfer score (area under the receiver operating curve, 0.65), or Modified Early Warning Score (area under the receiver operating curve, 0.58; P value < 0.0001 for all comparisons). At a specificity of 95%, the derived model had a sensitivity of 28% compared with 15% for Stability and Workload Index for Transfer score and 7% for the Modified Early Warning Score. Accuracy improvements with the derived model over Modified Early Warning Score and Stability and Workload Index for Transfer were similar in the Medical Information Mart for Intensive Care-III cohort.

Conclusions: A machine learning approach to predicting intensive care unit readmission was significantly more accurate than previously published algorithms in both our internal validation and the Medical Information Mart for Intensive Care-III cohort. Implementation of this approach could target patients who may benefit from additional time in the intensive care unit or more frequent monitoring after transfer to the hospital ward.

Keywords: intensive care unit; machine learning; patient readmission; quality improvement; prediction score

Determining who is ready for intensive care unit (ICU) discharge is a daily challenge for any ICU care team. Patients who experience unplanned readmissions to the ICU have increased mortality, length of stay, and cost compared with those not readmitted during their hospital stay (1, 2). Many physicians rely on clinical intuition and local hospital policies on the level of care that can be provided on the ward to determine who is ready to leave the ICU. This decision also is often influenced by the number of available ICU beds (3). The combination of external pressure to act as a steward of an often scarce hospital resource and variability of clinician intuition results in dramatic variation in the severity of illness of patients transferred out to the wards across ICUs (3, 4). Previous studies that performed retrospective reviews of ICU readmissions suggest that more than 10% of these readmissions are potentially preventable (5–7). For these reasons, improving risk stratification for patients at risk of clinical deterioration on the hospital ward after ICU discharge could have important benefits for critically ill hospitalized patients.

There has been growing interest in developing clinical decision support tools because of increasing availability of patient-level data (4, 8–15). Clinical decision support tools using real-time patient data may provide clinicians with additional information to guide decision making regarding timing of ICU discharge and level of monitoring needed on the ward. However, the tools developed to date have used only a fraction of the large amount of data available to clinicians with modern electronic health records. In addition, few of these tools have been validated, and the validation studies performed to date have demonstrated poor accuracy (8, 16). New complex classification algorithms have leveraged big data to create accurate predictive models for human behavior and events (17). These algorithms, known as machine learning, are flexible modeling techniques designed to both learn and generalize from data. Recently, published work has shown that several machine learning methods are more accurate than logistic regression models and early warning scores for predicting clinical deterioration in hospitalized patients (18–20). Therefore, a model created using a machine learning method may improve the prediction accuracy of an ICU readmission score. The aim of this study was to develop a novel

prediction tool on the basis of a machine learning algorithm called gradient boosted machine (GBM), using variables that are available in real time in the electronic health record, and to compare the derived model to previously published risk scores.

Methods

Setting and Study Population

We conducted an observational cohort study of all adult patients who received care in an ICU and were subsequently transferred to the medical-surgical wards at the University of Chicago Medical Center between November 1, 2008, and January 15, 2016. The University of Chicago Institutional Review Board approved the study protocol and granted a waiver of consent on the basis of minimal level of harm and general impracticability (IRB #16995). External validation of the derived model was performed in the Medical Information Mart for Intensive Care (MIMIC-III) database (version 1.4), which is a clinical database consisting of more than 38,000 ICU patients (medical, surgical, coronary care, and neonatal) admitted to Beth Israel Deaconess Medical Center (Boston, MA) from June 2001 to October 2012 (21, 22). The establishment of the database was approved by the Institutional Review Boards of the Massachusetts Institute of Technology (Cambridge, MA) and Beth Israel Deaconess Medical Center (Boston, MA). Our access to the database was approved by its administrators. The MIMIC-III cohort used in our analysis was composed of all adult patients who received care in an ICU and were subsequently transferred to the medical-surgical wards.

Outcomes

The primary outcome of interest was ICU readmission, defined as a transfer from the ICU to the wards and then back to the ICU during the same hospitalization. Patients who suffered a ward cardiac arrest (i.e., loss of a pulse with attempted resuscitation) after ICU-to-ward transfer were also classified as an ICU readmission. ICU readmissions were determined using location-stamped vital signs, and death on the wards was confirmed using administrative databases. Patients who died during the ward stay after ICU discharge without being readmitted or suffering a cardiac arrest were considered to be comfort care patients and were excluded

from the analysis. In addition, patients who were discharged to hospice after transfer to the ward were excluded from the analysis. Finally, patients readmitted to the ICU after an interventional or surgical procedure were not included in the readmitted cohort, as these readmissions could have been planned by the clinical service before the procedure.

Predictor Variables

Predictors were selected *a priori* on the basis of our clinical experience and previous literature describing risk factors for ICU readmission (2, 4, 8, 10, 12, 14, 15, 23, 24). These included demographic variables (i.e., age, sex, body mass index), vital signs, routinely collected laboratory values, medications administered during the ICU admission (i.e., vasopressors, sedatives, antibiotics, diuretics), ICU interventions (i.e., red blood cell transfusion orders, dialysis, mechanical ventilation), nursing documentation (i.e., Braden scale, Morse score), diagnostic tests (i.e., chest X-ray, computed tomography scan, electrocardiogram), and *International Classification of Diseases, Ninth Revision* codes from prior admissions at the same hospital (*see* Table E1 in the online supplement). The data for this study were provided by the Clinical Research Data Warehouse maintained by the Center for Research Informatics at University of Chicago and the Massachusetts Institute of Technology Lab for Computational Physiology (22). For each predictor variable, the measured value closest to the time of ICU discharge was used. In addition, vital sign trends and laboratory trends over the last 24 hours before ICU discharge were included as predictor variables, as these trends have been shown to improve the accuracy in predictions of impending clinical deterioration (25). Descriptive statistics were used to characterize the patient population and distributions of these predictor variables, and *t* tests, Wilcoxon rank sum tests, and chi-square tests were used, as appropriate, to compare characteristics for patients readmitted to the ICU to those who were not.

Model Development

To develop and then compare the prediction model to previous algorithms, the patient population was divided into derivation (60%) and internal validation (40%) cohorts by random number generator. The decision was made *a priori* to use a decision-tree–based

modeling technique, GBM, which is a collection of decision trees that are increasingly focused on the hard-to-predict cases to increase the predictive accuracy of classification problems. GBM is an ensemble algorithm, where each decision tree is trained one after another. Each subsequent decision tree is trained primarily with data that had been incorrectly classified by previous decision trees. This allows the model to gradually focus more on difficult-to-predict cases. This modeling technique has been shown to be one of the most accurate and best calibrated machine learning methods for predicting clinical deterioration on the wards (18). A GBM model was fit to develop the risk prediction model in the training cohort using the last ICU observation before ward transfer using all predictor variables. In addition, a simpler physiologic GBM model was fit that included age, vital signs, and laboratory data as predictor variables in the University of Chicago cohort to be validated in the MIMIC-III cohort. Because tree-based machine learning models have decreased accuracy in highly unbalanced data (e.g., many more nonevent observations than event observations), each patient who experienced

a readmission was matched to a randomly selected patient who did not experience a readmission to create a balanced dataset for model derivation (26, 27). The optimal number of splits for each individual tree, the total number of trees, and shrinkage factor were determined using tenfold cross-validation in the training dataset.

Accuracy Comparisons

Predicted probabilities were calculated for each ICU discharge in the validation dataset using the optimally tuned GBM model from the derivation dataset. The accuracy of the GBM model was compared with the previously published Stability and Workload Index for Transfer (SWIFT) score (4) and the Modified Early Warning Score (MEWS) (9, 28) in both the internal and MIMIC-III validation cohorts using sensitivity, specificity, and area under the receiver operating curve (AUC) for ICU readmission at any time after ICU discharge (29). The AUCs were compared according to the method of DeLong and colleagues (30). In addition, secondary accuracy comparisons were performed by classifying readmissions as early (≤72 hours after ICU discharge) or late (>72 hours after ICU

discharge) (31). A *post hoc* sensitivity analysis was performed, where those patients who died on the ward with no cardiac arrest and those who were discharged from the hospital to hospice were included in the analysis as nonevents to ensure model stability. To inspect the clinical validity of the derived model, a variable importance measure that used the change in the Gini index was used to visualize the contribution of the predictor variables in the most accurate model (27). The effects of the most accurate predictor variables across different values were also created for the derived model using partial dependence plots (32). All analyses were performed using R version 3.3.0 (The R Foundation for Statistical Computing) and Stata version 14.1 (StataCorp). A P value < 0.05 denoted statistical significance.

Results

During the study period, a total of 22,936 patients received care in an ICU in the University of Chicago cohort. These patients had 32,661 unique ICU admissions. Of the 32,661 ICU admissions, 7,776 (24%) ICU
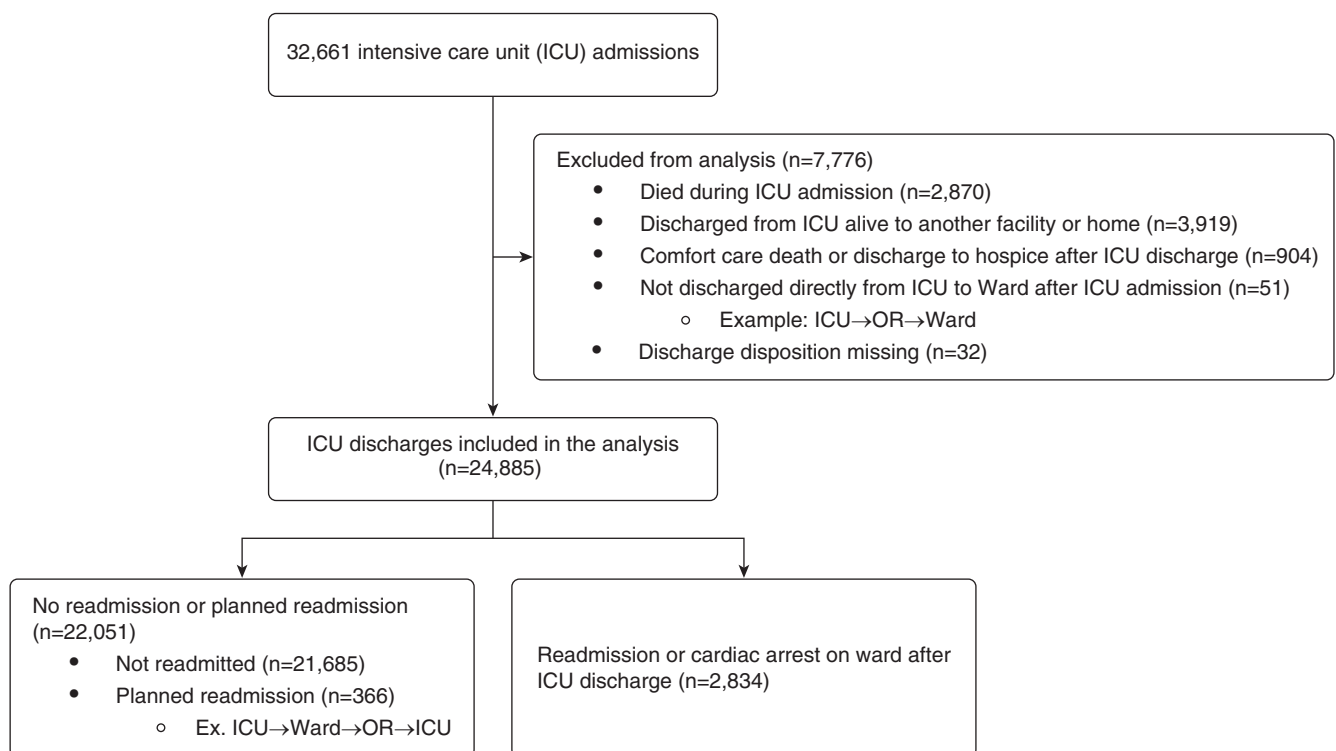
Figure 1. Patient flow diagram, including outcomes and patients excluded from analysis. ICU = intensive care unit; OR = operating room.

admissions were excluded from analysis (Figure 1), leaving a total of 18,000 patients with 24,885 ICU admissions that were followed by discharge to the ward included in the analysis. Eleven percent (2,834) of the ICU admissions that were discharged to the ward were readmitted to the ICU during the same hospitalization. The median time to readmission was 65 hours. Comparison of baseline demographics by ICU readmission status for entire cohort is shown in Table 1. Readmitted patients were more likely to be older (mean age, 61 yr [standard deviation (SD), 15 yr] vs. 58 yr [SD, 16 yr]; $P <$ 0.0001), male (54% vs. 51%; $P =$ 0.010), and underweight (9% vs. 6%; $P <$ 0.001). In addition, readmitted patients had a longer ICU length of stay than those who were not readmitted (3.9 d vs. 2.9 d; $P <$ 0.0001). Readmitted patients were also significantly more likely to have received several interventions and medications during their prior ICU stay, including mechanical ventilation (36% vs. 29%; $P <$ 0.0001), noninvasive ventilation (10% vs. 7%; $P <$ 0.0001), dialysis (16% vs. 7%; $P <$ 0.0001), vasopressors (20% vs. 13%; $P <$ 0.0001), intravenous sedatives (28% vs. 21%; $P <$ 0.0001), and red blood cell transfusions

(13% vs. 7%; $P <$ 0.0001; Table E2). The MIMIC-III external validation cohort included 42,303 ICU admissions that were followed by discharge to the ward. These patients had an observed ICU readmission rate of 8.2% (3,458 of 42,303).

In the internal validation dataset, the derived model better predicted ICU readmission at all time points after transfer to the ward than the MEWS and SWIFT scores (Table 2). The machine learning–derived model had the highest AUC for predicting those patients ever readmitted (AUC, 0.76; 95% confidence interval [CI], 0.75–0.78), followed by SWIFT (AUC, 0.65; 95% CI, 0.63–0.66), and MEWS (AUC, 0.58; 95% CI, 0.56–0.60) in the interval validation cohort. In the MIMIC-III cohort, the simpler physiologic machine learning derived in the University of Chicago cohort had the highest AUC for predicting those patients ever readmitted (AUC, 0.71; 95% CI, 0.70–0.72), followed by SWIFT (AUC, 0.58; 95% CI, 0.57–0.59), and MEWS (AUC, 0.57; 95% CI, 0.56–0.58). These machine learning–derived models were more accurate than the SWIFT score and MEWS in both the internal and MIMIC-III validation cohorts ($P <$ 0.001).

Table 2. Area under the receiver operating characteristic curve comparisons at different time points for readmission in the internal validation cohort

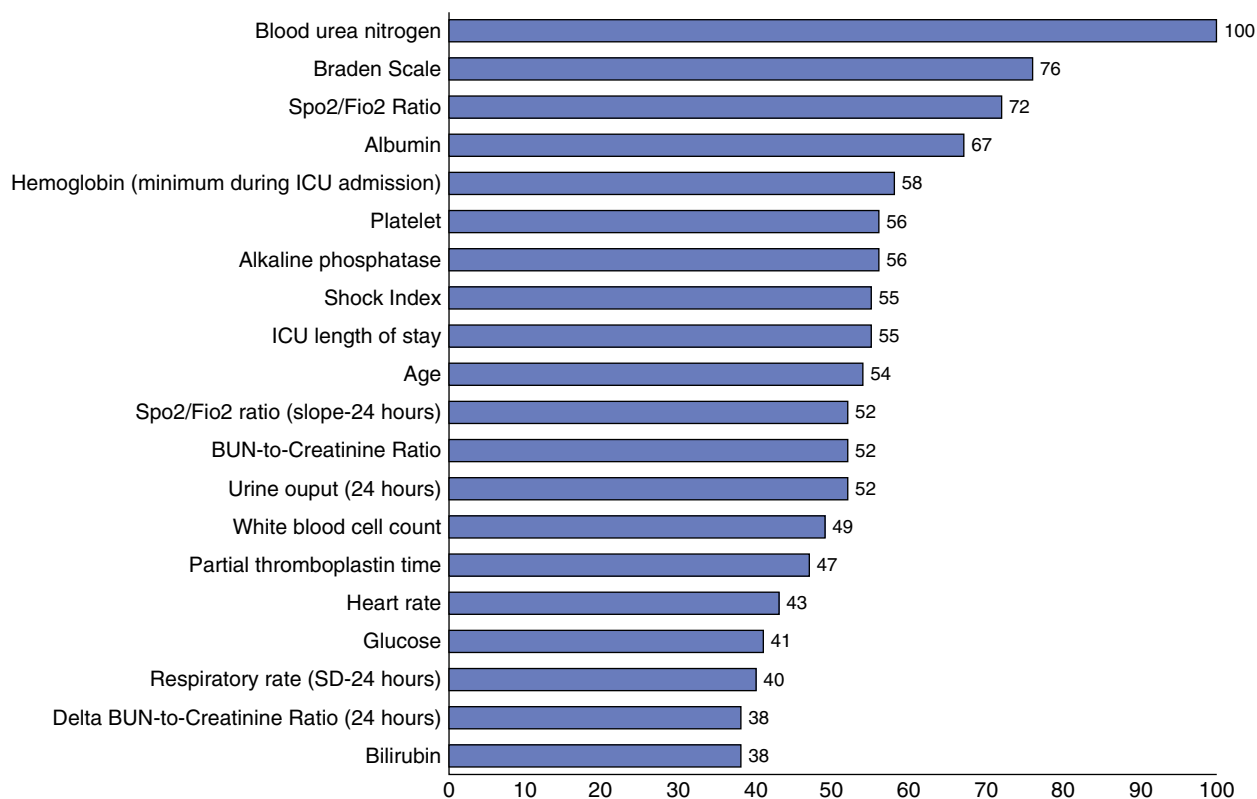| Time to Readmission	Model Type	AUC	95% CI
Early*	GBM	0.73	0.71–0.75
	SWIFT	0.62	0.60–0.65
	MEWS	0.60	0.58–0.62
Late†	GBM	0.77	0.75–0.78
	SWIFT	0.65	0.63–0.68
	MEWS	0.55	0.52–0.57
Ever	GBM	0.76	0.75–0.78
	SWIFT	0.65	0.63–0.66
	MEWS	0.58	0.56–0.60

Definition of abbreviations: AUC = area under the receiver operating characteristic curve; CI = confidence interval; GBM = gradient boosted machine; MEWS = Modified Early Warning Score; SWIFT = Stability and Workload Index for Transfer.
*Early indicates any ICU readmission that occurred between 0 and 72 hours after ICU discharge.
†Late indicates any ICU readmission that occurred more than 72 hours after ICU discharge.

In the University of Chicago internal validation cohort, a *post hoc* sensitivity analysis including ICU discharges followed by death on the hospital ward without a cardiac arrest or discharge to hospice had an AUC of 0.75. Furthermore, at a specificity of 95%, the derived model had a sensitivity of 28% compared with 15% for SWIFT score and 7% for the MEWS (Table E3). The addition of nursing, medication, and ICU intervention variables improved prediction of ICU readmission compared with physiologic parameters alone (AUC, 0.74 for physiologic variables vs. AUC, 0.76 for full model; $P <$ 0.0001; Table E4).

Blood urea nitrogen, Braden scale, oxygen saturation/fraction of inspired oxygen ratio, and albumin were the most important predictor variables in the full GBM model (Figure 2). The partial plots illustrating the effects of these predictors across a range of values in the model are shown in Figure 3. As shown, the risk for the readmission was increased at lower values of the Braden scale and oxygen saturation/ fraction of inspired oxygen ratio. The risk of readmission is highest for patients with Braden scale lower than 13. Risk also increased with increasing values of blood urea nitrogen, and an inflection point with more rapidly increasing risk occurred

Table 1. Comparison of baseline demographics by intensive care unit readmission status in the University of Chicago cohort

	Readmitted	Not Readmitted
Patients (N = 24,885)	2,834 (11)	22,051 (89)
Age, yr, mean (SD)	61 (15)	58 (16)
Sex, female	1,312 (46)	10,772 (49)
Race		
White	1,194 (42)	9,058 (41)
African American	1,359 (48)	11,130 (50)
Other/unknown	281 (10)	1,863 (9)
Body mass index, kg/m^2		
<18.5	247 (9)	1,369 (6)
18.5–24.9	938 (33)	6,973 (32)
25–29.9	780 (28)	6,143 (28)
30–39.9	425 (15)	3,733 (17)
≥40	444 (16)	3,833 (17)
Night discharge	1,101 (39)	7,807 (35)
Initial hospital location		
ER	522 (19)	4,469 (20)
Ward	920 (32)	4,564 (21)
OR	407 (14)	5,700 (26)
Other*	985 (35)	7,318 (33)
ICU length of stay, d, mean (SD)	3.9 (4.9)	2.9 (3.7)
In-hospital mortality	753 (26)	31 (0)

Definition of abbreviations: ER = emergency room; ICU = intensive care unit; OR = operating room; SD = standard deviation.
Data presented as n (%) unless otherwise noted.
*Other prior location included patients transferred from interventional areas or outside hospital transfers directly to the intensive care unit.

Figure 2. Twenty most important predictor variables in the gradient boosted machine model, scaled to a maximum of 100. The scaled numbers represent the relative importance of each variable, which is calculated by weighing model improvement as a result of each time a variable is used to split the data, averaged over all trees in the final model. BUN = blood urea nitrogen; $F_{I_{O_2}}$ = fraction of inspired oxygen; ICU = intensive care unit; SD = standard deviation; Sp_{O_2} = oxygen saturation as measured by pulse oximetry.

at blood urea nitrogen value of around 35 mg/dl. Last, the risk of readmission is highest for patient with albumin lower than 2.7 g/dl.

Discussion

Using data available in real time in the electronic health record, we have developed and validated a machine learning model to predict the likelihood that patients will return to the intensive care unit after transfer out to the wards. This is the first use of this novel machine learning technique to model intensive care unit readmission in the medical literature. Our model was more accurate than the Stability and Workload Index for Transfer score and the Modified Early Warning Score in both our interval validation and the Medical Information Mart for Intensive Care-III cohort. This model could be used to improve the allocation of scarce and expensive resources for patients transferred out of the intensive care unit. Specifically, the use of our model

could identify high-risk patients who could benefit from closer monitoring on the wards after transfer, including rapid response team visitation, and enhanced intensive care unit to ward handoffs.

The goal of our study was to create a risk score that could be calculated in real time from data in the electronic health record using a novel machine learning technique. Ensemble machine learning models, such as gradient boosted machine, have been shown to be more accurate than conventional logistic regression to classify disease or predict clinical outcomes in a variety of clinical settings (18, 33, 34). These machine learning approaches can account for nonlinear and higher dimensional relationships between a multitude of variables that enhance both prediction and explanatory power (35). To date, a machine learning model has not been developed to predict intensive care unit readmission. However, machine learning models have been shown to be more accurate than prior models using logistic regression in predicting hospital readmission in patients

with congestive heart failure (36). In addition to our novel modeling technique, we aimed to use variables available in real time in the electronic health record. Several previously published risk scores included variables that may not be accurately documented in real time, including admission diagnosis category or type of shock on admission to intensive care unit (8, 10, 12). These variables were not available in our dataset and are not abstracted reliably from the electronic health record.

The accuracy findings for Modified Early Warning Score and Stability and Workload Index for Transfer were similar to those of other studies validating these scores as prediction tools for intensive care unit readmission (16, 37, 38). The Stability and Workload Index for Transfer Stability score appears to lose predictive accuracy when externally validated on a mixed surgical and medical intensive care unit population (16). In addition, Rosa and colleagues also showed that the Stability and Workload Index for Transfer score had the same
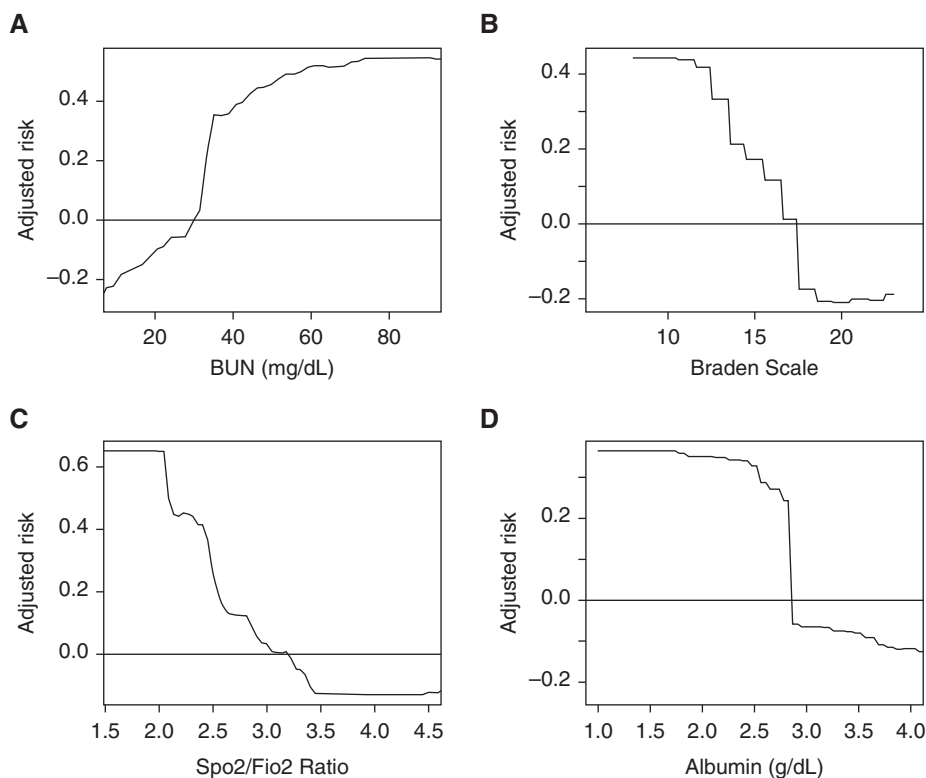
Figure 3. Partial plot of the effect of (A) blood urea nitrogen (BUN), (B) Braden scale, (C) oxygen saturation as measured by pulse oximetry/fraction of inspired oxygen (Sp_{O_2}/Fi_{O_2}) ratio, and (D) albumin on the risk of readmission to the intensive care unit across different values in the gradient boosted machine model.

predictive accuracy as tools such as the Sequential Organ Failure Assessment score and simplified Therapeutic Intervention Scoring System created and designed for different purposes (37). These prior studies, in combination with our findings, reinforce that prior prediction tools such as the Stability and Workload Index for Transfer score may not provide much value above clinical judgment. Our model, which uses a novel machine learning technique and included physiologic variables developed on a large mixed medical and surgical intensive care unit patient population, did not lose significant predictive accuracy when externally validated in the Medical Information Mart for Intensive Care-III cohort.

Previous studies developing models for intensive care unit readmission have varied widely in the outcomes investigated. Prior models have included readmissions within 48 hours of intensive care unit discharge, within 7 days of discharge, and anytime during the hospitalization (4, 8–15). Because different investigators have used differing outcome time points and because of the lack

of consensus on timeline for preventable intensive care unit readmission, we illustrated model AUCs for early and late readmission to illustrate the association between readmission timing and accuracy. As shown, our model was the most accurate for both early and late readmission. Interestingly, both our model and the Stability and Workload Index for Transfer score demonstrated improved accuracy for late readmissions compared with early readmissions. In addition, there has been wide variation regarding how death on the wards after discharge from the intensive care unit has been incorporated in these analyses. For example, some groups included death without intensive care unit readmission in the primary outcome, and others have excluded these patients (8). We chose to exclude patients who died on the wards without cardiac arrest or intensive care unit readmission, because these patients likely did not desire care escalation (e.g., were comfort care). Thus, they are not relevant to investigating risk factors for intensive care unit readmission, because they would not be readmitted despite being critically

ill. A *post hoc* sensitivity analysis including these comfort care deaths as readmissions did not significantly alter the accuracy of our model.

We investigated prior literature to determine the risk factors often associated with intensive care unit readmission and incorporated these as variables in our model. In light of the prior literature, it was not surprising that some of the most important predictor variables in the model were physiologic variables (4, 39). Specifically, one of the most important variables in our model was the oxygen saturation as measured by pulse oximetry/fraction of inspired oxygen ratio. This ratio has been shown to be a reliable noninvasive alternative to the arterial oxygen pressure/ fraction of inspired oxygen ratio (40). The arterial oxygen pressure/fraction of inspired oxygen ratio was included as a predictor in the most widely studied and used intensive care unit readmission score, the Stability and Workload Index for Transfer score (4). Interestingly, one of the most important predictors in our model was the Braden Scale. There has been increasing evidence that nursing assessments are correlated with in-hospital and postdischarge mortality (41). Furthermore, tools such as the Braden Scale and the Morse fall risk score have been shown to accurately predict 90-day mortality and discharge to an extended care facility (42). These readily available standardized nursing assessments appear to function well as a surrogate measure of patient frailty. Similar to other studies, we found that the inclusion of nonphysiologic variables (e.g., medications, intensive care unit interventions) in addition to physiologic variables improved model accuracy (8).

We should also note limitations of the current study. As our risk score was developed using patient data from a single institution, the generalizability of our findings will need further validation. However, our model continued to perform well in the Medical Information Mart for Intensive Care-III cohort. Important factors such as intensive care unit bed availability and hospital census can impact intensive care unit readmissions and, therefore, the prediction of these events. We did not have these important hospital-level variables in our model. Finally, we did not have information on reasons for intensive care unit admission. Because of this lack of data, we could not confidently exclude all patients who had elective admissions to the intensive

care unit. However, only a small proportion of our sample was likely to have planned intensive care unit admission from areas outside of the operating room or procedural areas, thereby having little impact on our risk score.

Conclusions

In summary, we developed and validated an intensive care unit readmission prediction model using a novel machine learning modeling technique. This risk score was more accurate than previously published tools and includes variables available in real time in the electronic health record. Our algorithm could be inserted into an electronic health record, a dashboard, or even a mobile app that could be used by both intensive care unit teams and rapid response teams to design targeted interventions aimed at reducing morbidity and mortality for patients transferred out of the intensive care unit. ■

References

1 Kramer AA, Higgins TL, Zimmerman JE. Intensive care unit readmissions in U.S. hospitals: patient characteristics, risk factors, and outcomes. *Crit Care Med* 2012;40:3–10.

2 Rosenberg AL, Watts C. Patients readmitted to ICUs : a systematic review of risk factors and outcomes. *Chest* 2000;118:492–502.

3 Skowronski GA. Bed rationing and allocation in the intensive care unit. *Curr Opin Crit Care* 2001;7:480–484.

4 Gajic O, Malinchoc M, Comfere TB, Harris MR, Achouiti A, Yilmaz M, et al. The Stability and Workload Index for Transfer score predicts unplanned intensive care unit patient readmission: initial development and validation. *Crit Care Med* 2008;36:676–682.

5 Durbin CG Jr, Kopel RF. A case-control study of patients readmitted to the intensive care unit. *Crit Care Med* 1993;21:1547–1553.

6 Snow N, Bergin KT, Horrigan TP. Readmission of patients to the surgical intensive care unit: patient profiles and possibilities for prevention. *Crit Care Med* 1985;13:961–964.

7 Al-Jaghbeer MJ, Tekwani SS, Gunn SR, Kahn JM. Incidence and etiology of potentially preventable ICU Readmissions. *Crit Care Med* 2016;44:1704–1709.

8 Hosein FS, Bobrovitz N, Berthelot S, Zygun D, Ghali WA, Stelfox HT. A systematic review of tools for predicting severe adverse events following patient discharge from intensive care units. *Crit Care* 2013;17:R102.

9 Reini K, Fredrikson M, Oscarsson A. The prognostic value of the Modified Early Warning Score in critically ill patients: a prospective, observational study. *Eur J Anaesthesiol* 2012;29:152–157.

10 Badawi O, Breslow MJ. Readmissions and death after ICU discharge: development and validation of two predictive models. *PLoS One* 2012;7:e48758.

11 Fernandez R, Baigorri F, Navarro G, Artigas A. A modified McCabe score for stratification of patients after intensive care unit discharge: the Sabadell score. *Crit Care* 2006;10:R179.

12 Ouanes I, Schwebel C, Français A, Bruel C, Philippart F, Vesin A, et al. A model to predict short-term death or readmission after intensive care unit discharge. *J Crit Care* 2012;27:422.e1–9.

13 Fernandez R, Serrano JM, Umaran I, Abizanda R, Carrillo A, Lopez-Pueyo MJ, et al.; Sabadell Score Study Group. Ward mortality after ICU discharge: a multicenter validation of the Sabadell score. *Intensive Care Med* 2010;36:1196–1201.

14 Daly K, Beale R, Chang RW. Reduction in mortality after inappropriate early discharge from intensive care unit: logistic regression triage model. *BMJ* 2001;322:1274–1276.

15 Frost SA, Tam V, Alexandrou E, Hunt L, Salamonson Y, Davidson PM, et al. Readmission to intensive care: development of a nomogram for individualising risk. *Crit Care Resusc* 2010;12:83–89.

16 Kastrup M, Powollik R, Balzer F, Röber S, Ahlborn R, von Dossow-Hanfstingl V, et al. Predictive ability of the stability and workload index for transfer score to predict unplanned readmissions after ICU discharge. *Crit Care Med* 2013;41:1608–1615.

17 Iwashyna TJ, Liu V. What's so different about big data? A primer for clinicians trained to think epidemiologically. *Ann Am Thorac Soc* 2014;11:1130–1135.

18 Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit Care Med* 2016;44:368–374.

19 Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of clinical deterioration in hospitalized adult patients with hematologic malignancies using a neural network model. *PLoS One* 2016;11:e0161401.

20 Allyn J, Allou N, Augustin P, Philip I, Martinet O, Belghiti M, et al. A comparison of a machine learning model with EuroSCORE II in predicting mortality after elective cardiac surgery: a decision curve analysis. *PLoS One* 2017;12:e0169772.

21 Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.

22 Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000;101:E215–E220.

23 Metnitz PG, Fieux F, Jordan B, Lang T, Moreno R, Le Gall JR. Critically ill patients readmitted to intensive care units–lessons to learn? *Intensive Care Med* 2003;29:241–248.

24 Rosenberg AL, Hofer TP, Hayward RA, Strachan C, Watts CM. Who bounces back? Physiologic and other predictors of intensive care unit readmission. *Crit Care Med* 2001;29:511–518.

25 Churpek MM, Adhikari R, Edelson DP. The value of vital sign trends for detecting clinical deterioration on the wards. *Resuscitation* 2016;102:1–5.

26 He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 2009;21:1263–1284.

27 Kuhn M, Johnson K. Applied predictive modeling. New York: Springer; 2013.

28 Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM* 2001;94:521–526.

29 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839–843.

30 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–845.

31 Ho KM, Dobb GJ, Lee KY, Finn J, Knuiman M, Webb SA. The effect of comorbidities on risk of intensive care readmission during the same hospitalization: a linked data cohort study. *J Crit Care* 2009;24:101–107.

32 Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001;29:1189–1232.

33 Weiss JC, Page D, Peissig PL, Natarajan S, McCarty C. Statistical relational learning to predict primary myocardial infarction from electronic health records. *Proc Innov Appl Artif Intell Conf* 2012;2012:2341–2347.

34 Dodd S, Berk M, Kelin K, Zhang Q, Eriksson E, Deberdt W, et al. Application of the Gradient Boosted method in randomised clinical trials: participant variables that contribute to depression treatment efficacy of duloxetine, SSRIs or placebo. *J Affect Disord* 2014;168:284–293.

35 Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure

hospitalization in patients with heart failure. *JACC Heart Fail* 2014;2: 429–436.

36 Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li SX, *et al*. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes* 2016;9: 629–640.

37 Rosa RG, Roehrig C, Oliveira RP, Maccari JG, Antonio AC, Castro Pde S, *et al*. Comparison of unplanned intensive care unit readmission scores: a prospective cohort study. *PLoS One* 2015;10:e0143127.

38 Rojas JC, Lyons PG, McCauley L, Picart J, Snyder AM, Kilaru M, *et al*. Accuracy Of Clinician Intuition And Comparison To An Early Warning Score For Predicting Intensive Care Unit Readmission [abstract]. *Am J Respir Crit Care Med* 2017;195:A6815.

39 Kaben A, Correa F, Reinhart K, Settmacher U, Gummert J, Kalff R, *et al*. Readmission to a surgical intensive care unit: incidence, outcome and risk factors. *Crit Care* 2008;12:R123.

40 Bilan N, Dastranji A, Ghalehgolab Behbahani A. Comparison of the Spo2/Fio2 ratio and the Pao2/Fio2 ratio in patients with acute lung injury or acute respiratory distress syndrome. *J Cardiovasc Thorac Res* 2015;7:28–31.

41 Rothman MJ, Solinger AB, Rothman SI, Finlay GD. Clinical implications and validity of nursing assessments: a longitudinal measure of patient condition from analysis of the Electronic Medical Record. *BMJ Open* 2012;2:e000646.

42 Tapper EB, Finkelstein D, Mittleman MA, Piatkowski G, Lai M. Standard assessments of frailty are validated predictors of mortality in hospitalized patients with cirrhosis. *Hepatology* 2015;62:584–590.