

RESEARCH ARTICLE

Open Access



Hospital acquired pressure injury prediction in surgical critical care patients

Jenny Alderden^{1*} , Kathryn P. Drake², Andrew Wilson³, Jonathan Dimas¹, Mollie R. Cummins¹ and Tracey L. Yap⁴

Abstract

Background: Hospital-acquired pressure injuries (HAPris) are areas of damage to the skin occurring among 5–10% of surgical intensive care unit (ICU) patients. HAPris are mostly preventable; however, prevention may require measures not feasible for every patient because of the cost or intensity of nursing care. Therefore, recommended standards of practice include HAPris risk assessment at routine intervals. However, no HAPris risk-prediction tools demonstrate adequate predictive validity in the ICU population. The purpose of the current study was to develop and compare models predicting HAPris among surgical ICU patients using electronic health record (EHR) data.

Methods: In this retrospective cohort study, we obtained data for patients admitted to the surgical ICU or cardiovascular surgical ICU between 2014 and 2018 via query of our institution's EHR. We developed predictive models utilizing three sets of variables: (1) variables obtained during routine care + the Braden Scale (a pressure-injury risk-assessment scale); (2) routine care only; and (3) a parsimonious set of five routine-care variables chosen based on availability from an EHR and data warehouse perspective. Aiming to select the best model for predicting HAPris, we split each data set into standard 80:20 train:test sets and applied five classification algorithms. We performed this process on each of the three data sets, evaluating model performance based on continuous performance on the receiver operating characteristic curve and the F_1 score.

Results: Among 5,101 patients included in analysis, 333 (6.5%) developed a HAPris. F_1 scores of the five classification algorithms proved to be a valuable evaluation metric for model performance considering the class imbalance. Models developed with the parsimonious data set had comparable F_1 scores to those developed with the larger set of predictor variables.

Conclusions: Results from this study show the feasibility of using EHR data for accurately predicting HAPris and that good performance can be found with a small group of easily accessible predictor variables. Future study is needed to test the models in an external sample.

Keywords: Pressure ulcer/pressure injury, Risk assessment, Hospital-acquired condition

Background

Hospital-acquired pressure injuries (HAPris), formerly called pressure ulcers or bedsores, are areas of injury to the skin and underlying tissue caused by external

pressure, usually over a bony area. In the United States, costs attributable to HAPris exceed \$26 billion a year [1]. HAPris are considered mostly preventable and defined as a "never event" or "serious reportable event" by the National Quality Forum [2]. HAPris occur among 5–10% of critical-care patients [3], with the highest risk among surgical critical-care patients [4].

Most HAPris are preventable; however, prevention may require using measures not feasible for every patient

*Correspondence: Jenny.Alderden@utah.edu

¹ University of Utah College of Nursing, 10 S 2000 E, Salt Lake City, UT 84112, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

because of cost or nursing time [5]. Therefore, recommended standards of practice include pressure injury (PrI) risk assessment at each nursing shift and with changes in patient status [6]. Unfortunately, the risk-assessment tools currently in use, such as the widely used Braden Scale [7], classify most critical-care patients as "high risk" [8–13], and therefore do not give critical-care nurses the information they need to allocate limited PrI prevention resources appropriately. A PrI risk-assessment tool allowing nurses to differentiate PrI risk among critical-care patients is imperative.

Electronic health records (EHRs) and data analytics can improve HAPrI risk assessment. Recent advances in machine learning (ML) present an opportunity to modernize and enhance future HAPrI risk assessment using the extensive data readily available in the EHR. Risk assessment and predicting future events are areas where combining modern ML techniques may identify novel patterns unapparent to humans to predict a target (in our case, HAPrI development). The benefits of ML approaches are particularly relevant in the ICU setting because of the dynamic physiologic nature of critical care patients. Unlike traditional prognostic tools like the Braden Scale, an ML approach can incorporate non-linear, complex interactions among variables (including correlated variables) [14].

The purpose of the current study was to develop and compare models predicting HAPrIs among surgical critical-care patients using EHR data. The specific aims were to (1) develop and compare predictive models, and (2) develop and compare more parsimonious predictive models using data readily available—and easily accessible—in the EHR.

Methods

Design

We chose a retrospective cohort design. All data were entered into structured fields in the EHR (EPIC[®]) and then obtained via a query from our institution's enterprise data warehouse (EDW).

Sample

Adult patients (aged ≥ 18 years) who were admitted to the surgical intensive care unit (SICU) or cardiovascular surgical intensive care unit (CVICU) at a Level 1 trauma center and academic medical center in the western United States between 2014 and 2018 were included in the sample. We included patients with community-acquired (present on admission) PrIs because those patients have risk for subsequent HAPrIs [15]. We excluded patients admitted to the ICU for less than 24 h because these short-stay patients were unlikely to manifest a HAPrI with this duration of exposure [16]. In those

patients with multiple ICU stays, our data were limited to the first ICU stay. At our facility, all ICU patients are placed on pressure redistribution mattresses [17].

Measures

HAPrI outcomes

There are six stages of HAPrIs defined by the National Pressure Injury Advisory Panel (NPIAP) [18]. Stage 1 HAPrIs are areas of nonblanching redness or discoloration in intact skin. Stage 2 HAPrIs represent partial-thickness tissue loss with exposed, viable dermis. Stage 3 HAPrIs are full-thickness wounds not extending into muscle, bone, or tendon. Stage 4 HAPrIs are full-thickness wounds extending down to muscle, tendon, or bone. Deep-tissue injuries (DTIs) are areas of intact or nonintact skin with a localized area of persistent, nonblanchable, deep-red, maroon, or purple discoloration revealing a dark wound bed or blood-filled blister. Finally, unstageable HAPrIs are areas of full-thickness tissue loss. These cannot be evaluated because eschar or slough obscures the area. Among these stages, only Stage 1 HAPrIs are reversible, sometimes resolving within hours after pressure offloading.

The outcome variable was the development of a HAPrI (Stages 2–4, DTI, or unstageable injury); Stage 1 PrIs were not included because they are often reversible, sometimes healing within hours [19, 20]. We used the initial (first recorded) HAPrI for patients with multiple injuries. Nurses at our hospital conduct a head-to-toe skin inspection once per shift, recording information about HAPrIs in a structured field in the EHR (EPIC[®]). We used the structured fields from the nursing documentation section of the EHR to record HAPrIs. We chose to use the structured fields from nursing documentation because a prior study showed that International Classification of Diseases codes underreport HAPrIs [21], likely because HAPrIs are noticed by nurses during head-to-toe skin assessments but not always flagged for inclusion in physician notes or billing codes. Furthermore, we relied on nursing documentation because it is the only accurate source of data for the date and time the HAPrI developed. All possible HAPrIs recorded in the nursing documentation at our facility are verified by a certified wound nurse, thus ensuring accuracy.

Potential predictor variables

Using Coleman and colleagues' conceptual framework for PrI etiology [22], we conducted a systematic review of the literature aimed at identifying risk factors for HAPrIs among critical-care patients [3]. Additionally, we conducted interviews with subject-matter experts, including clinicians at our study site: wound nurses, ICU nurses, intensivist physicians, surgeons, dieticians,

physical therapists, nursing assistants, anesthesiologists, and other healthcare team members. The result of this formative work is reflected in Table 1.

We developed our predictive models utilizing three different sets of variables, depicted in Fig. 1. The first set included all of the available data: data produced in the EHR during routine care and Braden Scale scores recorded by nurses. The second set of predictor variables included only data produced during routine care (excluding Braden Scale scores). The third data set had a parsimonious group of five variables we selected because they were easily accessible in EHRs from a data-warehouse perspective.

Analysis

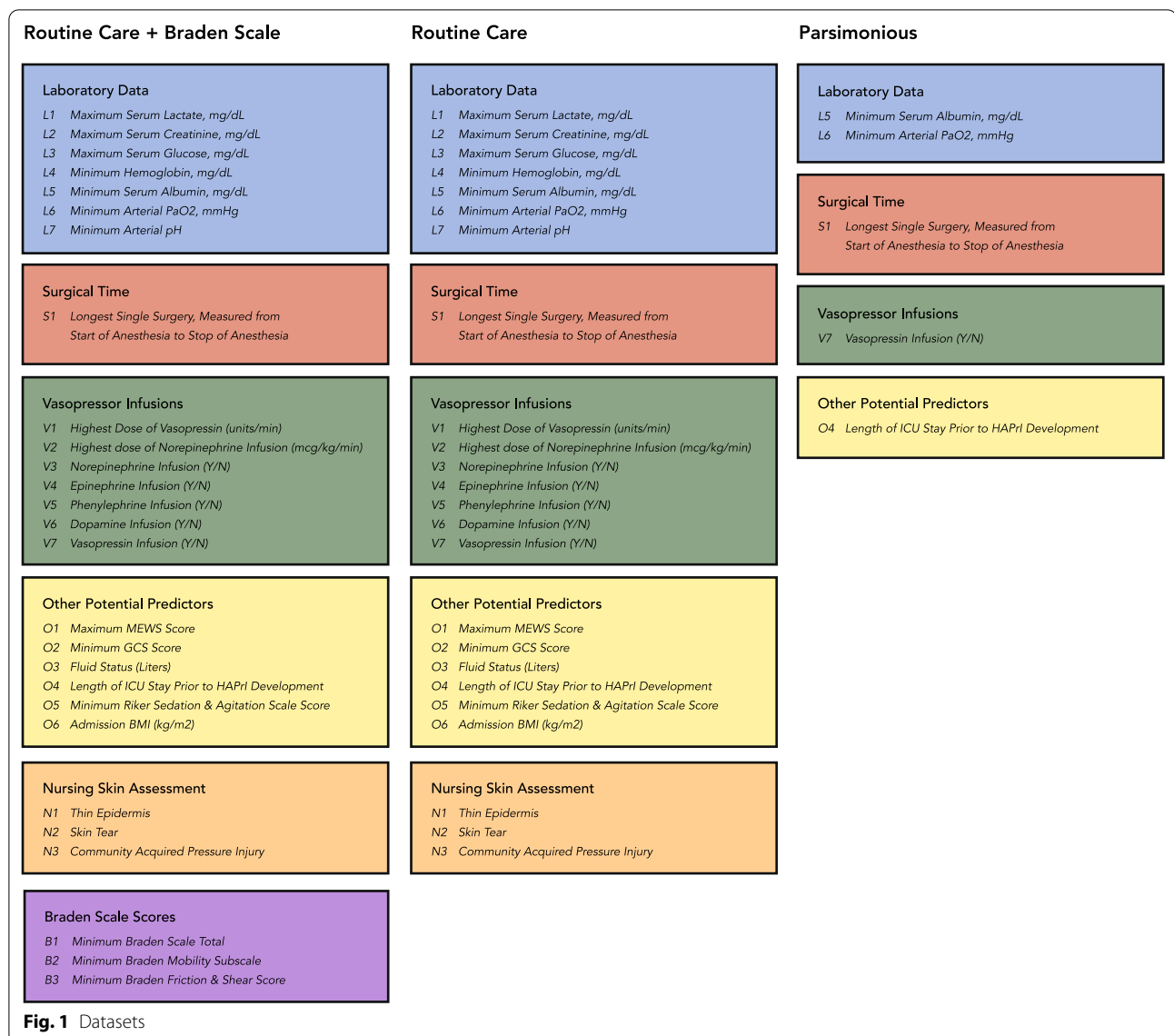
Data preprocessing

We obtained data from our study site's EDW; the query was performed by a team of biomedical informaticists. We limited our data to events occurring before the HAPrI outcome to mitigate target leakage (also called data leakage). Target leakage happens when a predictive model includes information unavailable when the model is applied to prospective data.

By validating the data in the human-readable system, we ensured that the query was accurate in terms of variable definitions and meaning and date/time stamps. A practicing critical-care nurse who worked within the EPIC® EHR system compared information obtained in

Table 1 Predictor variables

Laboratory data		
L ₁	Serum lactate	Maximum serum lactate (mg/dL)
L ₂	Serum creatinine	Maximum serum creatinine (mg/dL)
L ₃	Serum glucose	Maximum serum glucose (mg/dL)
L ₄	Hemoglobin	Minimum hemoglobin (mg/dL)
L ₅	Serum albumin	Minimum serum albumin (mg/dL)
L ₆	Arterial PaO ₂	Minimum arterial partial pressure of oxygen (mmHg)
L ₇	Arterial pH	Minimum arterial (pH)
Nursing skin assessment data		
N ₁	Thin epidermis	Thin epidermis with subcutaneous tissue loss
N ₂	Skin tear	Presence of a skin tear
N ₃	Community acquired pressure injury	Community-acquired pressure injury present at admission
Surgical time		
S ₁	Longest surgery	Longest single surgery, measured from start of anesthesia to stop of anesthesia
Vasopressor infusions		
V ₁	Vasopressin dose	Highest dose of vasopressin (units/min)
V ₂	Norepinephrine dose	Highest dose of norepinephrine infusion (mcg/kg/min)
V ₃	Norepinephrine infusion	Norepinephrine received (yes/no)
V ₄	Epinephrine infusion	Epinephrine received (yes/no)
V ₅	Phenylephrine infusion	Phenylephrine received (yes/no)
V ₆	Dopamine infusion	Dopamine received (yes/no)
V ₇	Vasopressin infusion	Vasopressin received (yes/no)
Other potential predictors		
O ₁	MEWS score	Maximum Modified Early Warning score
O ₂	GCS	Minimum Glasgow Coma Scale score
O ₃	Fluid status	Maximum daily net intake (in L)
O ₄	Length of ICU stay prior to HAPrI	Length of stay in the ICU prior to HAPrI development
O ₅	Riker score	Minimum Riker Sedation and Agitation Scale Score
O ₆	BMI	Admission Body Mass Index (kg/m ²)
Braden Scale Scores		
B ₁	Braden Scale total	Minimum total Braden Scale Score
B ₂	Braden Mobility Subscale	Minimum Braden Mobility Subscale Score
B ₃	Braden Friction and Shear	Minimum Braden Friction and Shear Subscale Score



the query against data in the human-readable system for 30 randomly selected participants, including 15 participants with HAPrIs. We chose the number 30 because, in earlier work conducted within a legacy EHR system, it was necessary to review up to 10 cases to find some inconsistencies; we tripled that figure to be conservative [23].

Missing data

Missing data were quantified and then assessed for patterns of missingness using graphical clustering displays. For variables not informatively missing, we replaced the missing data with values found using single-value random forest imputation. For clinically informatively missing variables, known by clinicians to be associated with

increased severity of illness, we substituted a "0" for missing and a "1" for not missing.

Imbalanced classes

Due to the nature of the data, there is a strong class imbalance of HAPrIs; more specifically, the majority of patients in ICUs do not develop HAPrIs, and this can present a problem when applying ML methods for prediction. For example, in models designed to optimize accuracy, the resulting predictive algorithms can end up classifying all outcomes as belonging to the majority class, and the accuracy of the model will be considered high. To remedy the high-class imbalance, we employed synthetic minority oversampling technique (SMOTE) in the training data [24]. By applying SMOTE

to the data before the algorithms, we ensured that models would better classify the outcomes in the minority class.

In addition to sensitivity and specificity, we considered both precision (also called positive predictive value) and recall (also called sensitivity) in evaluating model performance. The precision and recall of a model focus on the proportion of true positives and are calculated by

$$precision = \frac{TPs}{TPs + FPs} \tag{1}$$

and

$$recall = \frac{TPs}{TPs + FNs} \tag{2}$$

In words, precision is the ratio of correctly predicted positives to the total predicted positives, whereas recall is the ratio of correctly predicted positives to the total actual positives. The F_1 score is a summary score accounting for both precision and recall, defined as

$$F_1\text{Score} = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{3}$$

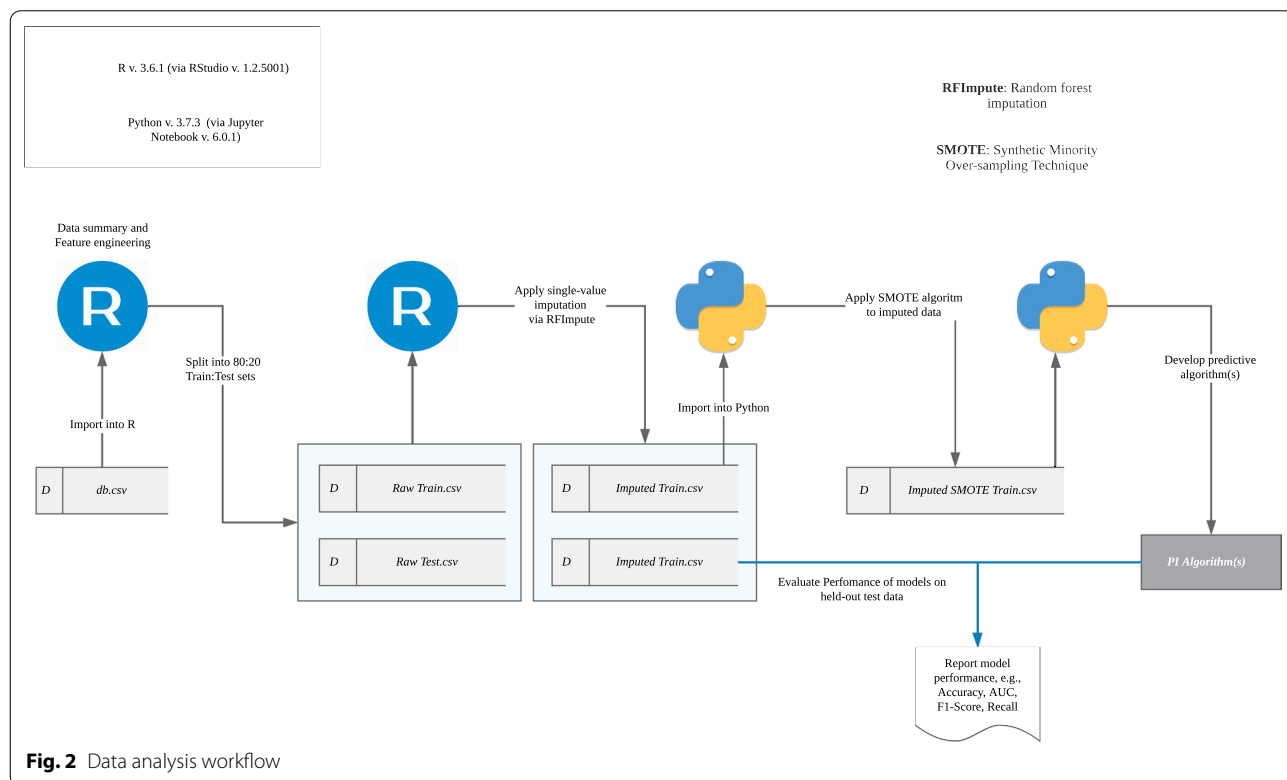
This quantity is the harmonic mean of precision and recall and preferable over a simple average because it penalizes extreme values.

All post-EPIC[®] data preparation (frequencies, assessment of missingness, imputation) was performed in R version 3.6.1. The SMOTE procedure and model development and evaluation were performed using Python version 3.7.0. Figure 2 depicts the data-analysis workflow.

Statistical analysis

To identify factors associated with HAPrIs, we conducted univariate analysis. We used a Mann–Whitney U test for ordinal variables, a *t*-test or Mann–Whitney U test for parametric versus nonparametric continuous variables, and a Pearson chi-square or Fisher exact test, as appropriate, for categorical variables.

For any given prediction or estimation problem, it is not possible beforehand to know the true form of relationships between variables and, therefore, not possible to know a priori which learning algorithm would be best. This can be particularly true when relationships between variables are complex, non-linear, and nested within each other. For this reason, our approach was to apply different types of models to different sets of variables (features) with different pre-processing techniques and compare the models' performance. The performance of each model was evaluated both as a binary classifier using a



confusion matrix and having continuous predicted probabilities for true positive and negative cases plotted as receiver operator characteristic (ROC) curves, with estimates of area under such curves (AUC). No post-prediction calibration was performed for generated models, although distributions of predicted probabilities were assessed (and probabilities from logistic regression are already naturally calibrated). Candidate models included keras (neural networks, with two hidden layers, each with 20 nodes each, and one output layer with a single node), random forests (ensembles of decision trees), gradient and adaptive (AdaBoost) boosting (algorithm focuses more on the examples that previous [internal] 'weak learners' misclassified, i.e., focuses more on *hard-to-train cases*), and multivariable logistic regression. All data processing and analysis steps were performed in Python v. 3.7 with all machine learning models available from the `sklearn` library [25]. For reproducibility and transparency, all code and results are hosted on and accessible through our Open Science Framework (OSF) page, <https://tinyurl.com/OSF-Predicting-HAPrI>.

Results

Sample characteristics

Our final sample consisted of 5,101 adult SICU and CVICU patients. One patient was excluded due to the absence of a medical record number; without the medical record number, we could not access data for other variables. The sample was predominantly male ($n = 3302$, 65%) and White ($n = 4256$, 83%). The mean age was 58 years ($SD = 17$).

HAPrI outcome

HAPrIs developed among 333 patients (6.5%), which is consistent with our institution's quarterly prevalence surveys.

Predictor variables

Table 2 outlines the univariate relationships between predictor variables and HAPrI outcomes. However, individual variable significance was not used to guide model development. Variables sets were selected based on ontologies presented in Fig. 1.

Predictive models

Confusion matrices for categorical (binary) prediction for each model and each set of features were produced and are available through the online supplement. Additionally, prediction probabilities vs. true event status are presented for each model as receiver operating characteristic (ROC) curves, with summary area under the curve (AUC) scores. The ROC curves using the Routine Care + Braden Scale data set for all five algorithms

are presented in Fig. 3 (AUC range: 0.774–0.812); ROC curves using the Routine Care data set are shown in Fig. 4 (AUC range: 0.761–0.822); and ROC curves in the Parsimonious data set are presented in Fig. 5 (AUC range: 0.775–0.821). F_1 scores for each type of algorithm across data sets (Routine Care + Braden Scale; Routine Care; Parsimonious) are shown in Table 3.

Although performance measures are generated on continuous scales, model performance is often evaluated within classes (from poor to excellent) and within the context of model deployment within the decision support setting. In this setting, model performance was comparable across the different data sets (Routine Care + Braden Scale; Routine Care; Parsimonious) and types of algorithms: neural networks (Keras), random forest, gradient boosting, adaptive boosting (AdaBoost), and multivariable logistic regression. Overall, the best performing model in the Routine Care + Braden Scale data set based on the AUC was logistic regression (AUC = 0.81), while random forest had the highest F_1 score ($F_1 = 0.36$). In the Routine Care data set, gradient boost generated the best discrimination based on the AUC (AUC = 0.82) and random forest and logistic regression tied for the highest F_1 score ($F_1 = 0.30$). Finally, Keras had the highest AUC at 0.82 in the parsimonious data set, while gradient boost had the highest F_1 score at 0.35.

Discussion

In this study, using data from surgical ICU patients within a single hospital system, we developed and compared models to predict HAPrIs in three data sets: Routine Care plus the Braden Scale, Routine Care, and a parsimonious set of five easily accessible predictor variables. This work represents a first step toward developing a model implementable in the EHR to provide real-time HAPrI risk assessment. In 2019, a guideline published by the NPIAP and its sister organizations, the European Pressure Injury Advisory Panel and the Pan Pacific Pressure Injury Alliance, recommended an automated approach to HAPrI risk assessment using data readily available in the EHR [5]. Advantages to harnessing data produced during routine care include more comprehensive evaluation of predictive factors related to HAPrI risk but missing from the Braden Scale (e.g., variables related to perfusion [3], severity of illness [26], and surgical factors [26–29]), the ability to update the risk prediction at frequent intervals with changes in patient status [9], and significantly decreasing the amount of time clinicians spend on data entry—currently, nurses manually record Braden Scale scores once per shift.

Our study results highlight the importance of considering class imbalance when evaluating model performance. The various types of models in our study all

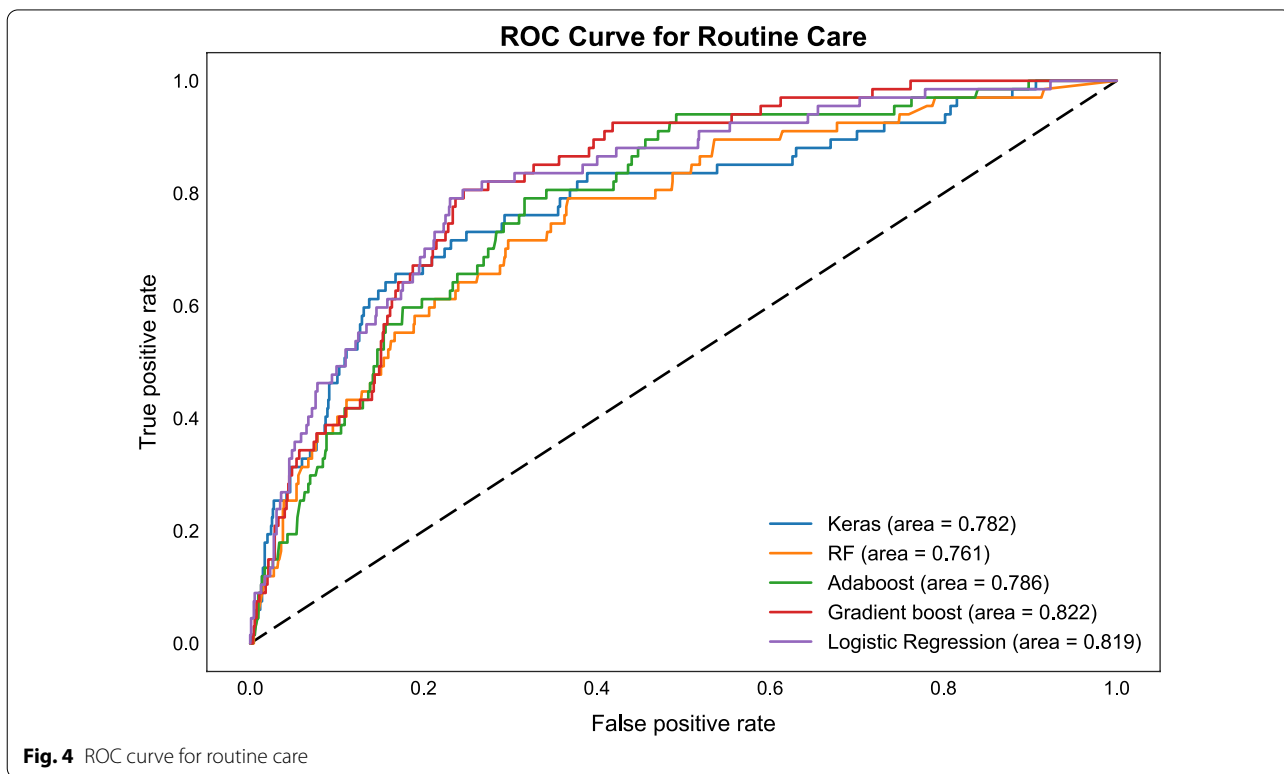
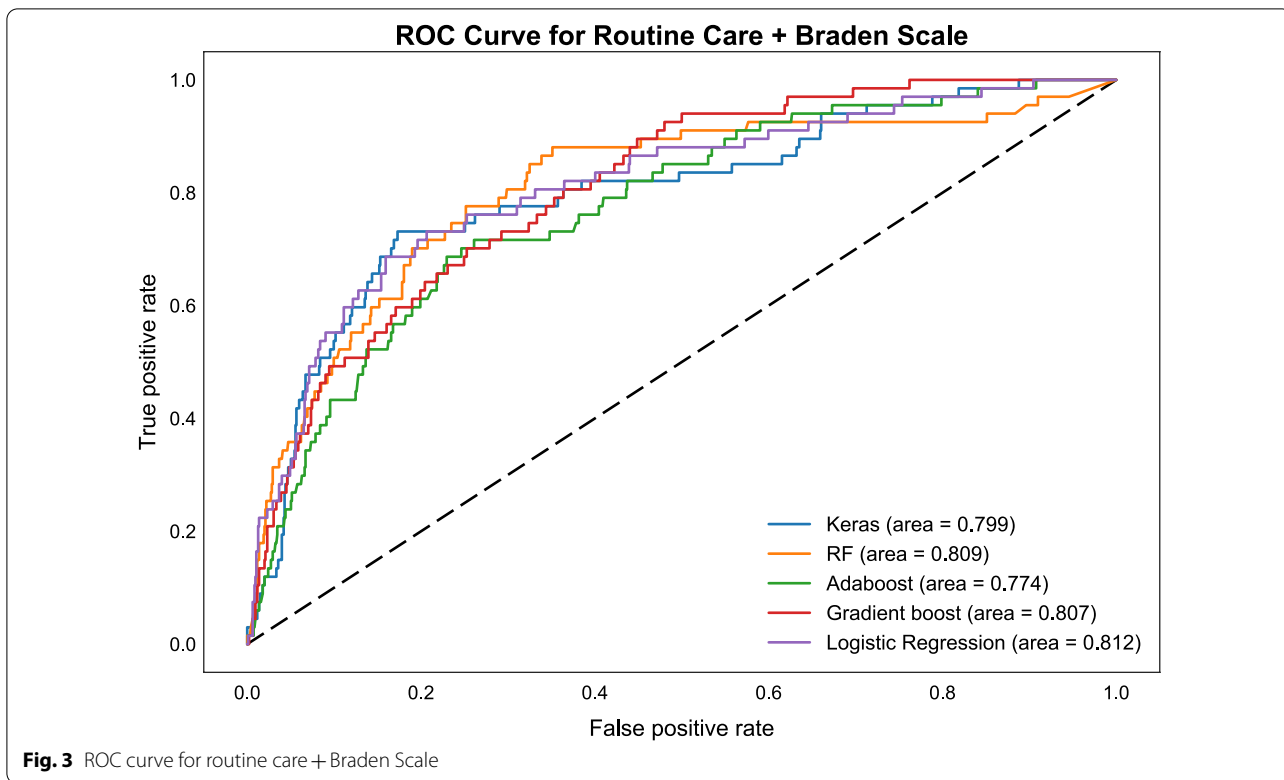
Table 2 Potential predictor variables and hospital-acquired pressure injury development

Variable	All patients	Patients with no HAPrI n = 4768	Patients with a HAPrI n = 333	p value
Demographic data				
Age in years, <i>M (SD)</i>	57.8 (17%)	58 (17%)	58 (15%)	<i>p</i> = .74
Sex, male, <i>n (%)</i>	3302 (65%)	3086 (65%)	216 (65%)	<i>p</i> = 1.0
Race, White, <i>n (%)</i>	4256 (83%)	3988 (84%)	268 (81%)	<i>p</i> = .13
Ethnicity, non-Hispanic, <i>n (%)</i>	4452 (87%)	4166 (87)	286 (86%)	<i>p</i> = .54
Laboratory data				
Serum lactate (mg/dL), <i>M (SD)</i>	4.0 (3.7)	3.9 (3.6)	5.9 (5.0)	<i>p</i> < .001
Serum creatinine (mg/dL), <i>M (SD)</i>	1.9 (1.9)	1.8 (1.9)	2.8 (2.1)	<i>p</i> < .001
Serum glucose (mg/dL), <i>M (SD)</i>	231 (148)	228 (143)	285 (200)	<i>p</i> < .001
Hemoglobin (mg/dL), <i>M (SD)</i>	9.0 (2.6)	9.1 (2.6)	7.6 (2.2)	<i>p</i> < .001
Serum albumin (mg/dL), <i>M (SD)</i>	3.1 (0.8)	3.2 (0.7)	2.7 (0.6)	<i>p</i> < .001
Arterial PaO ₂ (mmHg), <i>M (SD)</i>	54 (40)	55 (40)	46 (31)	<i>p</i> < .001
Arterial pH, <i>M (SD)</i>	7.27 (0.11)	7.27 (0.10)	7.22 (0.13)	<i>p</i> < .001
Nursing skin-assessment data				
Thin epidermis, <i>n (%)</i>	882 (18%)	801 (17%)	81 (24%)	<i>p</i> = .001
Skin tear, <i>n (%)</i>	641 (13%)	548 (11%)	93 (28%)	<i>p</i> < .001
Community-acquired pressure injury, <i>n (%)</i>	168 (3%)	138 (2.9%)	30 (9%)	<i>p</i> < .001
Surgical time				
Longest surgery, minutes, <i>M (SD)</i>	181 (156)	179 (152)	204 (197)	<i>p</i> < .001
Vasopressor infusions				
Vasopressin dose (units/min), <i>M (SD)</i>	0.01 (0.03)	0.01 (0.03)	0.02 (0.03)	<i>p</i> < .001
Norepinephrine dose (mcg/kg/min), <i>M (SD)</i>	0.06 (0.42)	0.05 (0.29)	0.25 (1.18)	<i>p</i> < .001
Norepinephrine (yes/no), <i>n (%)</i>	1154 (23%)	982 (21%)	172 (52%)	<i>p</i> < .001
Epinephrine (yes/no), <i>n (%)</i>	1490 (29%)	1352 (28%)	138 (41%)	<i>p</i> < .001
Phenylephrine (yes/no), <i>n (%)</i>	190 (4%)	173 (4%)	17 (5%)	<i>p</i> = .22
Dopamine (yes/no), <i>n (%)</i>	205 (4%)	186 (4%)	19 (5%)	<i>p</i> = .14
Vasopressin (yes/no), <i>n (%)</i>	1757 (34%)	1586 (33%)	171 (51%)	<i>p</i> < .001
Other potential predictors				
MEWS score, <i>M (SD)</i>	4.3 (1.9)	4.1 (2.0)	6.2 (1.9)	<i>p</i> < .001
GCS, <i>M (SD)</i>	9.2 (4.8)	9.3 (4.7)	6.8 (4.6)	<i>p</i> < .001
Fluid status (liters), <i>M (SD)</i>	2.3 (2.1)	2.2 (2.0)	3.7 (2.6)	<i>p</i> < .001
Length of ICU stay prior to HAPrI (days), <i>M (SD)</i>	5 (7)	4.6 (6.2)	13 (9.9)	<i>p</i> < .001
Riker score, <i>M (SD)</i>	2.8 (1.2)	2.9 (1.2)	2.1 (1.2)	<i>p</i> < .001
BMI (kg/m ²), <i>M (SD)</i>	30.1 (12.4)	30.1 (12.5)	30.7 (11.0)	<i>p</i> = .36
Braden Scale scores				
Braden Scale total score, <i>M (SD)</i>	13 (3)	13 (3)	11 (3)	<i>p</i> < .001
Braden Mobility subscale, <i>M (SD)</i>	2.0 (0.7)	2.1 (0.7)	1.6 (0.7)	<i>p</i> < .001
Braden Friction subscale, <i>M (SD)</i>	1.9 (0.5)	1.9 (0.5)	1.6 (0.6)	<i>p</i> < .001

HAPrI, hospital-acquired pressure injury; arterial PaO₂, arterial partial pressure of oxygen; MEWS, Modified Early Warning Scale; GCS, Glasgow Coma Scale; Riker, Riker Sedation and Agitation Scale; BMI, body mass index

demonstrated good performance on the ROC curve. However, because HAPrIs are relatively rare, occurring in < 10% of ICU patients, continuous performance on the ROC curve may overestimate a model's utility due to the large number of true negatives [30]. F_1 scores present an additional approach to evaluating the model's performance by utilizing a weighted average of

precision and recall. Employing F_1 scores was a strength in our study, enabling more accurate estimation of the model's true utility for predicting HAPrIs. For example, in the Routine Care + Braden Scale data set, gradient boosting was the best performer based on the area under the ROC curve, whereas logistic regression was the best performer based on F_1 score. However, the



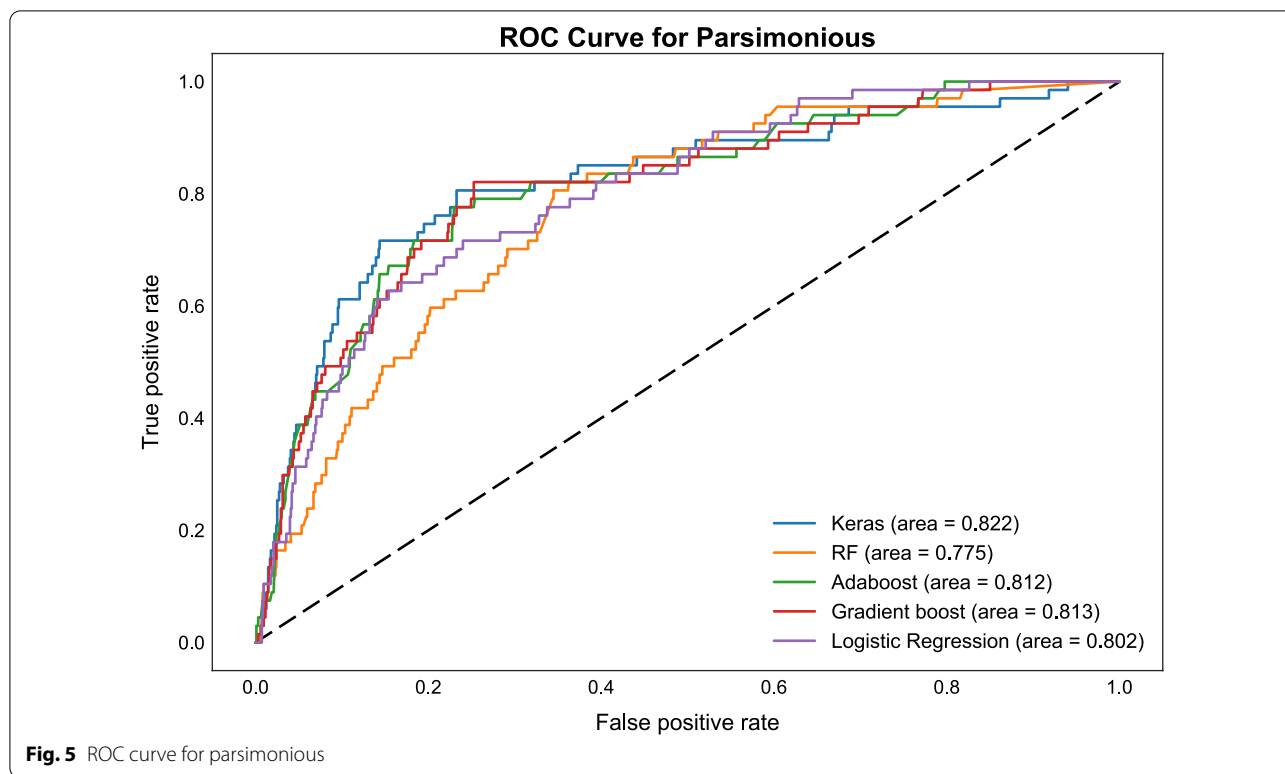


Table 3 F₁ scores for algorithms

	Keras	RF	AdaBoost	Gradient boost	Logistic regression
Routine care + Braden Scale	0.28	0.36	0.29	0.28	0.31
Routine care	0.27	0.30	0.26	0.26	0.30
Parsimonious	0.34	0.26	0.33	0.35	0.28

RF, random forest

gradient boosting algorithm was "accurate" because it primarily predicted the majority class—no HAPrI—and this is reflected in its lower F₁ score.

Beyond classifying variables as missing or observed before imputation, it may be worthwhile considering modeling the observation in future work; for example, using a Heckman two-stage modeling approach (a model for observation and a model for outcome) [31]. However, we did not pursue this as we were not as interested in effect estimation (and therefore not as concerned with bias in estimated coefficients) but in building light and transportable models generating accurate prediction scores.

In our study, the inclusion of the Braden Scale modestly improved the models' performance. In contrast, a

prior study found no improved performance when the Braden Scale was added to other variables obtained from routine-care documentation in the EHR [32]. Although including the Braden Scale did improve performance in our study, it requires clinicians to manually input six values during each shift. Reducing the time clinicians spend documenting is a worthy goal since documentation burden contributes to clinician burnout [33].

Models developed in the Parsimonious data set performed almost as well as the models developed using the larger number of predictor variables in the Routine Care data set. This finding is consistent with results from prior studies conducted among patients undergoing cardiovascular surgery [34, 35]. The Parsimonious model performance is encouraging in terms of clinical feasibility for eventual model implementation, demonstrating that a small group of informative variables can predict HAPrIs almost as well as a more comprehensive group of variables.

In our analysis, classic (parametric) multivariable logistic regression performed similarly to modern (non-parametric) machine learning models, including neural networks (Keras), random forest, gradient boosting, and adaptive boosting (AdaBoost). When performance is similar, logistic regression may be preferable to more complicated machine learning approaches for reasons of

interpretability, computing power, and compactness, i.e., logistic regression is human-interpretable (coefficients give an immediate indication of size and direction of effect) and less computationally expensive. Furthermore, logistic regression models can be incorporated and run within most EHR systems (including EPIC®), simplifying model implementation. Our finding that multivariable logistic regression showed good performance in predicting HAPrI compared to other machine learning methods is consistent with a previous study [32].

Our study includes several limitations. Although we sought to include a comprehensive set of predictor variables, it is possible that other unexamined variables would improve predictive performance. For example, we did not consider some potential risk factors, including medications such as corticosteroids [36] or events like intra-operative blood loss [27], because those factors were more challenging to obtain in our EHR. Similarly, we chose not to collect data about preventive interventions like repositioning because prior studies have shown that documentation in the EHR does not necessarily reflect what was 'really' done [37]—particularly relevant in the setting of a busy ICU, where nurses have many competing priorities. Preventive interventions, especially repositioning, are also confounded by critical illness because repositioning is sometimes not possible due to hemodynamic instability.

Conclusion

Overall, results from this study show the feasibility of using EHR data for accurately predicting HAPrIs and that good performance can be found with a small group of easily accessible predictor variables. Future study is needed to test the models in an external sample.

Abbreviations

CVICU: Cardiovascular surgical intensive care unit; EDW: Enterprise data warehouse; EHR: Electronic health record; HAPrI: Hospital-acquired pressure injury; ICU: Intensive care unit; ML: Machine learning; NPIAP: National Pressure Injury Advisory Panel; PrI: Pressure injury; ROC: Receiver operating characteristic; AUC: Area under the (ROC) curve; SICU: Surgical intensive care unit; SMOTE: Synthetic minority oversampling technique.

Acknowledgements

We are grateful for the ongoing input and assistance from the University of Utah Pressure Injury Prevention Committee, led by Ms. Kristy Honein.

Authors' contributions

All authors have read and approved the manuscript. JA, AW, and MC conceived the study idea and designed the study. KD and AW performed the analysis. KD developed and managed the Open Science Framework page. JA, AW, KD, JD, MC, and TY all contributed to data interpretation. JA, AW, KD, JD, MC, and TY co-wrote the manuscript. All read and approved the final manuscript.

Funding

This study was funded by the American Association of Critical-Care Nurses Critical Care Grant. This foundation grant provided funding for data extraction and analysis.

Availability of data and materials

The data used in this publication include protected health information, and therefore cannot be freely shared. Data sharing will be possible with case-by-case approval from our institution's Institutional Review Board; requests may be directed to the corresponding author.

Ethics approval and consent to participate

This study was approved by the University of Utah Hospital Institutional Review Board (IRB) (IRB#00111380). The IRB approved a waiver of informed consent per the United States Department of Health and Human Services guideline 21 CFR 46.

Consent for publication

All authors consent to publication.

Competing interests

The authors have no competing interests.

Author details

¹ University of Utah College of Nursing, 10 S 2000 E, Salt Lake City, UT 84112, USA. ² Department of Computer Science, Boise State University, 777 W Main Street, Boise, ID 83704, USA. ³ Parexel, 2520 Meridian Parkway, Durham, NC, USA. ⁴ Duke University School of Nursing, 307 Trent Drive, Durham, NC 27710, USA.

Received: 28 May 2020 Accepted: 13 December 2020

Published online: 06 January 2021

References

- Padula WV, Delarmente BA. The national cost of hospital-acquired pressure injuries in the United States. *Int Wound J*. 2019;16(3):634–40.
- Centers for Medicare & Medicaid Services. Eliminating serious, preventable, and costly medical errors—never events. 2006. <https://www.cms.gov/newsroom/fact-sheets/eliminating-serious-preventable-and-costly-medical-errors-never-events>, Accessed 16 Apr 2020.
- Alderden J, Rondinelli J, Pepper G, Cummins M, Whitney J. Risk factors for pressure injuries among critical care patients: a systematic review. *Int J Nurs Stud*. 2017;71:97–114.
- Chen HL, Chen XY, Wu J. The incidence of pressure ulcers in surgical patients of the last 5 years: a systematic review. *Wounds*. 2012;24(9):234–41.
- Emily Haesler, editor. European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel, and Pan Pacific Injury Alliance. Prevention and treatment of pressure ulcers/injuries: clinical practice guideline. The international guideline. EPUAP/NPIAP/PPPIA; 2019.
- Agency for Healthcare Research and Quality. Preventing pressure ulcers in hospitals. 2014. <https://www.ahrq.gov/professionals/systems/hospital/pressureulcer toolkit/putool1.html>. Accessed 12 Apr 2020.
- Bergstrom N, Braden BJ, Laguzza A, Holman V. The Braden Scale for predicting pressure sore risk. *Nurs Res*. 1987;36(4):205–10.
- Anthony D, Papanikolaou P, Parboteeah S, Saleh M. Do risk assessment scales for pressure ulcers work? *J Tissue Viability*. 2010;19(4):132–6.
- Black J. Pressure ulcer prevention and management: a dire need for good science. *Ann Intern Med*. 2015;162(5):387–8.
- Sharp CA, McLaws ML. Estimating the risk of pressure ulcer development: is it truly evidence based? *Int Wound J*. 2006;3(4):344–53.
- Cox J. Predictive power of the Braden Scale for pressure sore risk in adult critical care patients: a comprehensive review. *J Wound Ostomy Continence Nurs*. 2012;39(6):613–21 (quiz 22–3).
- Cho I, Noh M. Braden Scale: evaluation of clinical usefulness in an intensive care unit. *J Adv Nurs*. 2010;66(2):293–302.
- Hyun S, Vermillion B, Newton C, Fall M, Li X, Kaewprag P, et al. Predictive validity of the Braden Scale for patients in intensive care units. *Am J Crit Care*. 2013;22(6):514–20.
- Saria S, Butte A, Sheikh A. Better medicine through machine learning: what's real, and what's artificial? *PLoS Med*. 2018;15(12):e1002721.
- Coleman S, Gorecki C, Nelson EA, Closs SJ, Defloor T, Halfens R, et al. Patient risk factors for pressure ulcer development: systematic review. *Int J Nurs Stud*. 2013;50(7):974–1003.

16. Baumgarten M, Rich SE, Shardell MD, Hawkes WG, Margolis DJ, Langenberg P, et al. Care-related risk factors for hospital-acquired pressure ulcers in elderly adults with hip fracture. *J Am Geriatr Soc*. 2012;60(2):277–83.
17. Huang HY, Chen HL, Xu XJ. Pressure-redistribution surfaces for prevention of surgery-related pressure ulcers: a meta-analysis. *Ostomy Wound Manage*. 2013;59(4):36–48.
18. Edsberg LE, Black JM, Goldberg M, McNichol L, Moore L, Sieggreen M. Revised National Pressure Ulcer Advisory Panel Pressure Injury Staging System: revised pressure injury staging system. *J Wound Ostomy Cont*. 2016;43(6):585–97.
19. Halfens RJ, Bours GJ, Van Ast W. Relevance of the diagnosis “stage 1 pressure ulcer”: an empirical study of the clinical course of stage 1 ulcers in acute care and long-term care hospital populations. *J Clin Nurs*. 2001;10(6):748–57.
20. Alderden J, Zhao YL, Zhang Y, Thomas D, Butcher R, Zhang Y, et al. Outcomes associated with stage 1 pressure injuries: a retrospective cohort study. *Am J Crit Care*. 2018;27(6):471–6.
21. Ho C, Jiang J, Eastwood CA, Wong H, Weaver B, Quan H. Validation of two case definitions to identify pressure ulcers using hospital administrative data. *BMJ Open*. 2017;7(8):e016438.
22. Coleman S, Nixon J, Keen J, Wilson L, McGinnis E, Dealey C, et al. A new pressure ulcer conceptual framework. *J Adv Nurs*. 2014;70(10):2222–34.
23. Alderden J, Pepper GA, Wilson A, Whitney JD, Richardson S, Butcher R, et al. Predicting pressure injury in critical care patients: a machine-learning model. *Am J Crit Care*. 2018;27(6):461–8.
24. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*. 2013;14:106.
25. Scikit-learn: machine learning in Python. Scikit-learn.org. <https://scikit-learn.org/stable/index.html>. Accessed 17 Sept 2020.
26. Lima Serrano M, Gonzalez Mendez MI, Carrasco Cebollero FM, Lima Rodriguez JS. Risk factors for pressure ulcer development in intensive care units: a systematic review. *Med Intensiva*. 2017;41(6):339–46.
27. Chen HL, Jiang AG, Zhu B, Cai JY, Song YP. The risk factors of postoperative pressure ulcer after liver resection with long surgical duration: a retrospective study. *Wounds*. 2019;31(9):242–5.
28. Shen WQ, Chen HL, Xu YH, Zhang Q, Wu J. The relationship between length of surgery and the incidence of pressure ulcers in cardiovascular surgical patients: a retrospective study. *Adv Skin Wound Care*. 2015;28(10):444–50. <https://doi.org/10.1097/01.ASW.0000466365.90534.b0>.
29. O'Brien DD, Shanks AM, Talsma A, Brenner PS, Ramachandran SK. Intra-operative risk factors associated with postoperative pressure ulcers in critically ill patients: a retrospective observational study. *Crit Care Med*. 2014;42(1):40–7.
30. Saito T, Rehmsmeier M. The precision–recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*. 2015;10(3):e0118432.
31. Heckman JJ. Sample selection bias as a specification error. *Econometrica*. 1979;47(1):153. <https://doi.org/10.2307/1912352>.
32. Cramer EM, Seneviratne MG, Sharifi H, Ozturk A, Hernandez-Boussard T. Predicting the incidence of pressure ulcers in the intensive care unit using machine learning. *EGEMS (Wash DC)*. 2019;7(1):49.
33. Herndon RM. EHR is a main contributor to physician burnout. *J Miss State Med Assoc*. 2016;57(4):124.
34. Chen HL, Yu SJ, Xu Y, et al. Artificial neural network: a method for prediction of surgery-related pressure injury in cardiovascular surgical patients. *J Wound Ostomy Continence Nurs*. 2018;45(1):26–30. <https://doi.org/10.1097/WON.0000000000000388>.
35. Lu CX, Chen HL, Shen WQ, Feng LP. A new nomogram score for predicting surgery-related pressure ulcers in cardiovascular surgical patients. *Int Wound J*. 2017;14(1):226–32. <https://doi.org/10.1111/iwj.12593>.
36. Chen HL, Shen WQ, Xu YH, Zhang Q, Wu J. Perioperative corticosteroids administration as a risk factor for pressure ulcers in cardiovascular surgical patients: a retrospective study. *Int Wound J*. 2015;12(5):581–5. <https://doi.org/10.1111/iwj.12168>.
37. Yap TL, Kennerly SM, Simmons MR, et al. Multidimensional team-based intervention using musical cues to reduce odds of facility-acquired pressure ulcers in long-term care: a paired randomized intervention study. *J Am Geriatr Soc*. 2013;61(9):1552–9. <https://doi.org/10.1111/jgs.12422>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

