BMJ Open

# Predicting pressure injury risk in hospitalised patients using machine learning with electronic health records: a US multilevel cohort study

William V Padula ,[1,2,3] David G Armstrong ,[3,4] Peter J Pronovost ,[5,6] Suchi Saria [7,8]

**To cite:** Padula WV, Armstrong DG, Pronovost PJ, *et al*. Predicting pressure injury risk in hospitalised patients using machine learning with electronic health records: a US multilevel cohort study. *BMJ Open* 2024;**14**:e082540. doi:10.1136/bmjopen-2023-082540

► Prepublication history for this paper is available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-082540).

Received 26 November 2023
Accepted 06 March 2024

## ABSTRACT

**Objective** To predict the risk of hospital-acquired pressure injury using machine learning compared with standard care.

**Design** We obtained electronic health records (EHRs) to structure a multilevel cohort of hospitalised patients at risk for pressure injury and then calibrate a machine learning model to predict future pressure injury risk. Optimisation methods combined with multilevel logistic regression were used to develop a predictive algorithm of patient-specific shifts in risk over time. Machine learning methods were tested, including random forests, to identify predictive features for the algorithm. We reported the results of the regression approach as well as the area under the receiver operating characteristics (ROC) curve for predictive models.

**Setting** Hospitalised inpatients.

**Participants** EHRs of 35 001 hospitalisations over 5 years across 2 academic hospitals.

**Main outcome measure** Longitudinal shifts in pressure injury risk.

**Results** The predictive algorithm with features generated by machine learning achieved significantly improved prediction of pressure injury risk (p<0.001) with an area under the ROC curve of 0.72; whereas standard care only achieved an area under the ROC curve of 0.52. At a specificity of 0.50, the predictive algorithm achieved a sensitivity of 0.75.

**Conclusions** These data could help hospitals conserve resources within a critical period of patient vulnerability of hospital-acquired pressure injury which is not reimbursed by US Medicare; thus, conserving between 30 000 and 90 000 labour-hours per year in an average 500-bed hospital. Hospitals can use this predictive algorithm to initiate a quality improvement programme for pressure injury prevention and further customise the algorithm to patient-specific variation by facility.

## STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ We mined a cohort of over 35 000 hospitalisations from individual electronic health records to analyse changes in pressure injury risk over time.

⇒ Machine learning methods were used to a develop a logistic algorithm for predicting categorical shifts between low risk and high risk for pressure injury among hospitalised patients.

⇒ This machine learning algorithm outperforms standard of care (ie, risk assessment with the Braden Scale) to improve the prediction of patient risk for pressure injury.

⇒ This study provides an algorithm that can be applied to predict pressure injury risk in current patient populations, as well as a general methodology for developing in-house predictive algorithms for rare patient safety events.

⇒ A limitation of this study is inherent in biases of the Braden Scale towards Caucasian skin tones that may extend to this predictive algorithm based on its calibration data, unless future studies can cross-validate this algorithm with other races and ethnicities.

## INTRODUCTION

Pressure injuries are the fastest rising hospital-acquired condition according to the US Agency for Healthcare Research in Quality (AHRQ) and as a result have become the second most common reason for medical malpractice civil suits in the USA.[1] Pressure injuries (aka 'bedsores' or 'pressure ulcers') are harmful wounds that appear over bony prominences on the bodies of bedridden patients, usually caused by a combination of friction and shear.[2] The USA experiences approximately 60 000 deaths from pressure injuries among 2.5 million cases.[3] The total cost for US health systems to manage the acute needs of patients' pressure injuries during hospitalisation exceeds US$26 billion.[4] Long-term, chronic wound management caused by pressure injuries exceeds an additional US$20 billion in US Centers for Medicare and Medicaid Services (CMS) spending, tallying US$40–US$50 billion annually.[5]

To alleviate patient risk of pressure injury, health systems should invest in performance-improvement initiatives that increase



For numbered affiliations see end of article.

**Correspondence to**
Dr William V Padula; padula@usc.edu

compliance with routine tasks such as daily skin checks and risk assessment, as well as pressure offloading and turning every 2–4 hours.[6] However, compliance with these labour-intensive interventions challenges providers who must manage multiple priorities for varying patient types.[7] Perhaps this is why AHRQ identified a 6% increase in pressure injury rates between 2014 and 2017 in the USA, whereas rates of ten comparable hospital-acquired conditions (eg, infections, falls) have all decreased.[1]

Predictive analytics offers potential to alleviate some provider burden by automating latter iterations of the risk assessment process. Currently, providers must assess patient risk for pressure injury on admission and every 12–24 hours thereafter in acute care using a standardised, predictively valid instrument such as the Braden Scale.[8 9] Most electronic health record (EHR) systems collect information on risk, such as completed Braden Scales, along with other individualised outcome measures.[10] Braden scores of 18 or below dictate risk to the patient for pressure injury development, and subsequent follow-up with guideline-based practices. Using EHR data combined with these dichotomised risk scores could be used to predict the risk trajectories over a hospitalisation and save providers time on follow-up risk assessments of low-risk patients, thereby focusing on mitigating circumstances of pressure injury for select individuals.

Such 'early warning systems' of patient risk for various levels of deterioration in health, injury or infections that depend on machine learning are becoming more commonplace in healthcare, especially hospitals.[11] For instance, Modified Early Warning Score systems have been developed to trigger intervention in patients with general deteriorations in health that could increase risk of cardiac arrest.[12] In close alignment with pressure injury and other iatrogenic injuries, septic shock has received attention in a number of separate studies given the high risk of mortality associated with the onset of sepsis. Sepsis risk severity scoring systems such as the Acute Physiology and Chronic Health Evaluation II, Simplified Acute Physiology Score II, Sequential Organ Failure Assessment scores and Simple Clinical Score have been validated to assess illness severity and risk of death among septic patients.[13–16] However, more patient-specific scoring systems that adjust for individual health factors that could amplify the severity of septic shock have been innovated since then, such as the Targeted Real-Time Early Warning System.[17] One study attempted to use random forests to predict staged pressure injuries; however, predicting the outcome fails to actionise processes to mitigate patient risk based on specific risk factors (eg, mobility, nutrition).[18] Overall, these innovations offer insight to an early warning system for pressure injury prevention, which requires adjustment for individualised predictors of risk and remains operational during a patient's length of stay.

Using machine learning methods and a health system EHR of multiple contiguous facilities, we devised a predictive algorithm that could detect patients requiring risk mitigation by providers on high-risk trajectories for pressure injury. The model was trained on a cohort of hospitalised patients for extended durations of time and cross-validated with additional patient cases. As a result of this work, we present a predictive algorithm that could be integrated with EHRs in order detect patients on high-risk trajectories for pressure injury and prescribe follow-up with preventive care by bedside clinical staff.

## METHODS

### Study design

We applied supervised machine learning methods to an observational cohort of hospitalised patients in the USA to identify predictors of pressure injury risk and develop an algorithm to predict future risk. After multiple machine learning approaches were considered, we ultimately chose to develop an algorithm using multilevel modelling methods with predictors determined by random forests, in accordance with methodological guidelines prescribed by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) methodological task force on machine learning.[19] We also followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guideline for reporting on this study.[20] Multilevel modelling offered improved individual-specific prediction capabilities relative to other methods tested with the data at hand.[21] This approach also offered transparency to clinicians who may elect to use predictive analytics combined with clinical practice guidelines, such as validated risk assessment tools (eg, the Braden Scale), in order to make more informed decisions about patient care.[22]

### Data source

After receiving institutional review board (IRB) approval, we applied our method to the clinical data warehouses at two privately held US academic medical centre databases to collect samples of deidentified EHR data across four contiguous hospitals. An observational cohort of hospitalised patient EHRs over 5 years was used to obtain a substantial sample of calibration data for analytics from 2014 to 2019.

We identified 21 388 unique patients aged 18 years or older admitted to inpatient care at these facilities, including critical care, general medicine and surgery with at least 2 risk assessments for pressure injuries. The data were extracted from the Epic-based EHR system in use at these facilities. Data were gathered longitudinally by patient encounter every 12 hours during hospitalisation. Age and Braden risk score data for each 12-hour clinical shift up to the 12th shift (ie, 6 days total), and discharge diagnosis codes were collected at the patient- encounter level.

### Study population

EHRs were included in the study in accordance with CMS performance guidelines for hospital quality issued by the US Agency for Healthcare Research and Quality as

Patient-Safety Indicator #3 (PSI03, V.2020), which defines inclusion and exclusion criteria for hospital-acquired pressure injury.[23] A pressure injury is caused by damage to the skin and underlying tissue as a result of pressure and shearing forces; pressure injuries are measured in stages from 1 to 4, including unstageable, to describe advancement from low-stage pressure injury denoted by bruising and erythema, to high-stage pressure injuries that include open ulcerations that reach bone and muscle fascia.[24] PSI03 inclusion criteria for hospital-acquired pressure injury cases are as follows: stage 3, 4 or unstageable pressure injury not present-on-admission (not as a primary or secondary diagnosis), International Classification if Diseases Ninth-Revision (ICD-9, 707.x) and Tenth-Revision (ICD-10, L89.xx), per 1000 discharges among surgical and medical patients ages 18 years and older; excludes stays less than 3 days and certain diagnosis or procedural codes that place them at predisposed risk to skin injury: certain skin conditions (MDC 9); pregnancy (MDC 14); spinal cord injury (hemiplegia, paraplegia or quadriplegia); spina bifida; pedicled graft or debridement and a transfer between different facilities.

We modified some specifications of the PSI03 criteria to develop a predictive model that was inclusive of some high-risk patient populations, as consistent with prior informatics research on this patient population.[10] For instance, we included patients who were hospitalised for at least 5 days and obtained 2 or more risk scores during admission. We also included patients with spinal cord injury in the study cohort since providers could benefit from a greater understanding of the association between risk and pressure injury, based on previous findings linking these outcomes.[10]

### Risk assessment measure: the Braden Scale

The dataset contained Braden risk scores recorded by healthcare providers during every 12-hour clinical shift. The Braden Scale is an instrument that ranks risk on a scale from 6 to 23, with lower scores corresponding to higher pressure injury risk.[9] A score of 18 or less is considered the threshold for follow-up with routine prevention and recurring risk assessment according to international clinical practice guidelines.[8]

### Model development: logistic regression

To develop an algorithm for predicting an individual's risk of pressure injury development, we explored multiple supervised methods in machine learning. The Braden Scale provided dichotomous sets of risk, lending well to the development of a logistic regression model to formulate a decision rule on categorical shifts in risk from low to higher risk states. The logistic features of the machine learning model were meant to be prescriptive to clinicians by providing a decision rule of which patients required follow-up within 2–5 days window as necessary.

We explored fixed-effects logistic regression in comparison to multilevel logistic regression with random-effects (aka 'mixed-effects'). Mixed effects adjust models for subject-specific parameters, termed empirical Bayes estimates, in order to estimate changes in outcomes between individuals.[21]

Log-likelihood ratios (LLRs) were used in order to interpret the degree of fit of separate models. LLRs are based on a function of the error or variability in a particular model.[21] LLR tests can be conducted to estimate the degree of improved fit between models, therefore, we used $\chi^2$ tests of the LLRs to differentiate predictive performance between fixed-effects and random-effects models.[21] We also used LLR to classify the degree of enhanced predictive accuracy of models with additional data. By minimising the number of data points needed to provide an accurate estimate of future risk, we hoped to reduce the total cost of continuous risk assessment for non-critical patients.

### Model development: feature selection

To develop the predictive algorithm, the model necessitated features that captured patient-specific measurements of risk for pressure injury, as well as weighted estimates (ie, regression coefficients) using machine learning. Given the large sample size, we first applied established normalisation and regularisation techniques to the data structure to protect this analysis for risk of overfitting, in addition to cross-validation.[22] Features that would serve as predictors were obtained using a variety of supervised methods (ie, random forests, support vector machines, boosting, bagging, LASSO and Ridge regression).[25–27] After testing these methods, we determined that random forests provided relevant predictors at reduced mean-squared error compared with other approaches. Predictors for the algorithm were derived from random forests applied to randomly selected groups of 1000 observations at a time, and regressing 10 variables in each run. We simulated this process for 20 000 iterations and maintained predictors in the final model that remained statistically significant ($p<0.05$). In addition, we maintained the use of time and age—due to health system restrictions on deidentification, age was the only piece of obtainable demographic data. The covariate time was an expression of hospital length of- stay for each patient, broken down into 12-hour shifts, and age was a fixed effect exploring the impact of age on risk since the Braden Scale does not adjust for either factor. However, age has been explored in several studies to have a strong association with pressure injury risk.[7 10 28] These data were used to formulate a multilevel logistic regression model as framed below for predicting future risk (equation 1).

$$Logistic\ [risk\ score_{ij}] = (\beta_0 + \mu_{i0}) + (\beta_1 + \mu_{i1}) * time_{ij} + \beta_2 * \ age_{ij} + \cdots + \epsilon_{ij} \quad (1)$$

We tested the mixed effects form of the model for a random-intercept and random-slope focused on time. The empirical Bayes estimate for the intercept ($u_{i0}$) also accounted for within-subject correlation resulting from clustering encounters within patients, as it is hypothesised that patients with previous encounters resulting in a hospital-acquired pressure injury are greater risk for skin injury in a future encounter.[10 21]

For the multilevel approach, we tested for the presence of correlations between concurrent measurements within observations, such as the degree of collinearity between risk scores measured using the Braden Scale. Correlation between time point measures represents a legitimate concern because clinicians can be compelled to score a patient similarly to a previous score if it is available, increasing the degree of autocorrelation over time.[21] Various covariance structures were tested for in this approach, including independent, unstructured, autocorrelated and exponential. We also adjusted the quadrature of the models to achieve convergence. The final approach was configured to four quadrature points with independent covariance.

The regression approaches tested assumed linearity in subject-specific risk trajectories, and we applied quadratic random effects to samples to discern whether non-linearity could be observed.[21] The absence of non-linearity would rule out the need for more sophisticated (ie, deep learning) approaches to statistical prediction.

### Model evaluation and cross-validation

We applied constrained optimisation to the LLRs of each model to identify the optimal number of scores needed to efficiently and accurately predict future risk. By identifying the time point with the greatest marginal benefit to prediction accuracy, LLR, we identified the number of risk scores necessary for collection in a hospital setting to minimise overuse of costly skilled labour (ie, hospital staff) and resources.

Given the model's binary outcome measure and training on the dataset, we performed cross-validation to the analysis by organising test and training sets to evaluate the predictive validity of the model. Data were randomly separated at the encounter level, with 60% of the data used for training the model and 40% of the data used for the validation set. We then predicted the results of the regression model on the validation set.

We explored model accuracy using sensitivity and specificity to plot a receiver operating characteristics curve (ROC) and measure area under the ROC (AUROC). In this analysis, improved sensitivity represents an increase in the identification of patients at risk of a pressure injury, given that risk actually does exist (ie, true positive); specificity represents a reduction in the designation of at-risk patients who are not in fact considered at risk (ie, false positive). This balance in improved sensitivity and sensitivity leads to health system efficiency since clinicians would be providing follow-up care to patients who actually have a higher likelihood of pressure injury risk. The AUROC was compared with a limited time-adjusted and age-adjusted model, and with the performance of the Braden Scale alone (ie, standard care). The AUROC value would guide our understanding of specificity and sensitivity of the predictive algorithm.

### Financial impact

Pressure injury prevention is a costly protocol to implement on a daily basis. A typical hospitalised patient requires a routine that includes nursing time and material resources that can add up to US\$99.44 per patient per day on average. Thus, a large 500-bed hospital could spend up to US\$50 000 per day and US\$18 million per year on this standard practice. Based on the assumption that a predictive algorithm could provide prescriptive follow-up for a select group of about only 25% of a typical patient population, we used an existing economic model in order to calculate the budget impact that this algorithm could save hospitals.[29] We assumed that the budget impact was based on 100% compliance with an algorithm-driven protocol, adjusting for the sensitivity and specificity of the algorithm. These economic modelling assumptions are consistent with previous economic evaluations on the cost-effectiveness of pressure injury prevention.[29 30]

### Patient and public involvement

We engaged stakeholders in the design and execution of this research through the 'Pressure Ulcer Research

**Table 1** Study population

| Characteristic | N | Minimum | Median | Mean | Maximum |
|---|---|---|---|---|---|
| Patients | 21 388 | | | | |
| Encounters | 35 001 | | | | |
| Age | | 18.0 | 58.0 | 56.6 | 108.0 |
| Length of stay (days) | | 5.0 | 19.6 | 18.7 | 23.0 |
| Average Braden Score, by encounter | | 6.0 | 19.6 | 18.7 | 23.0 |
| Common clinical characteristics, by encounter | | | | | |
| Pressure injury, stage 3/4/unstageable (ICD-10 = L89.xx) | 1549 | | | | |
| Hyperlipidaemia (ICD-10 E78.xx) | 5656 | | | | |
| Spinal cord injury (ICD-10 S14.xx) | 99 | | | | |
| Bed confinement (ICD-9 Z74.01) | 49 | | | | |

ICD-9, International Classification of Diseases, Ninth-Revision; ICD-10, International Classification of Diseases, Tenth-Revision.

**Table 2** Pressure injury risk scores from multilevel regression models generated with predictors from a machine learning approach versus a restricted time-adjusted and age-adjusted model: time-adjusted and age-adjusted model

| Predictor | OR | Coefficient | SE | P value | (95% CI) | |
|---|---|---|---|---|---|---|
| Intercept | 0.017 | −4.077 | 0.001 | <0.001 | 0.014 | 0.020 |
| Time | 0.711 | −0.341 | 0.010 | <0.001 | 0.691 | 0.731 |
| Age | 1.051 | 0.501 | 0.001 | <0.001 | 1.049 | 1.054 |
| Variance (time) | 1.886 | 0.037 | | | 1.815 | 1.960 |
| Variance (intercept) | 6.587 | 0.112 | | | 6.371 | 6.809 |
| Log likelihood ratio | −69 095 | | | | | |

The key explanatory predictors of risk score in the full model were determined by applying random forests to electronic health records. Random effects were applied to a random-intercept and random slope on time.

**Table 3** Pressure injury risk scores from multilevel regression models generated with predictors from a machine learning approach versus a restricted time-adjusted and age-adjusted model: a random forest-generated model

| Predictor | OR | Coefficient | SE | P value | (95% CI) | |
|---|---|---|---|---|---|---|
| Intercept | 0.001 | −6.973 | 0.192 | <0.001 | −7.349 | −6.597 |
| Time | 1.194 | 0.177 | 0.014 | <0.001 | 0.151 | 0.204 |
| Age | 1.056 | 0.054 | 0.002 | <0.001 | 0.049 | 0.059 |
| Rx: beta-adrenergic and anticholinergic combinations | 3.479 | 1.247 | 0.092 | <0.001 | 1.067 | 1.427 |
| Rx: phosphate replacement | 1.210 | 0.191 | 0.088 | 0.031 | 0.018 | 0.364 |
| Rx: electrolyte replacement | 0.697 | −0.361 | 0.087 | <0.001 | −0.531 | −0.192 |
| Rx: erythropoetin stimulating agents | 3.987 | 1.383 | 0.200 | <0.001 | 0.990 | 1.776 |
| Rx: thiazide/diuretic | 0.769 | −0.263 | 0.139 | 0.059 | −0.535 | 0.010 |
| Rx: analgesic/Narcotic | 1.075 | 0.072 | 0.097 | 0.457 | −0.118 | 0.263 |
| Rx: zinc replacement | 2.597 | 0.954 | 0.207 | <0.001 | 0.549 | 1.359 |
| Rx: vasopressor | 7.097 | 1.960 | 0.255 | <0.001 | 1.461 | 2.459 |
| Dx: bed confinement (ICD-10 Z74.01) | 1857.170 | 7.527 | 1.045 | <0.001 | 5.478 | 9.576 |
| Dx: dyspnoea (ICD-10 R06.02) | 0.729 | −0.317 | 0.181 | 0.080 | −0.671 | 0.037 |
| Dx: hyperlipidaemia (ICD-10 E78.xx) | 0.745 | −0.295 | 0.094 | 0.002 | −0.479 | −0.110 |
| Dx: spinal cord injury (ICD-10 S14.xx) | 183 131.000 | 12.118 | 0.902 | <0.001 | 10.350 | 13.886 |
| Dx: chronic osteomyelitis (ICD-10 M86.xx) | 1749.917 | 7.467 | 0.833 | <0.001 | 5.835 | 9.100 |
| Order: upper extremity venous duplex exam | 4.241 | 1.445 | 0.474 | 0.002 | 0.516 | 2.374 |
| Lab: culture, stool | 0.336 | −1.092 | 0.181 | <0.001 | −1.447 | −0.737 |
| Lab: culture and stain, quantitative with Anaerobes | 4.856 | 1.580 | 0.273 | <0.001 | 1.045 | 2.115 |
| Lab: urinalysis chemistry screen | 4.485 | 1.501 | 0.084 | <0.001 | 1.336 | 1.665 |
| Lab: lipid panel | 2.095 | 0.740 | 0.113 | <0.001 | 0.518 | 0.962 |
| Lab: prealbumin panel | 1.841 | 0.610 | 0.102 | <0.001 | 0.411 | 0.809 |
| Lab: vancomycin random assay | 4.657 | 1.538 | 0.125 | <0.001 | 1.294 | 1.783 |
| Variance (time) | 0.441 | 0.015 | | | 0.413 | 0.471 |
| Variance (intercept) | 16.654 | 0.441 | | | 15.812 | 17.540 |
| Log likelihood ratio | −32 551 | | | | | |

The key explanatory predictors of risk score in the full model were determined by applying random forests to electronic health records. Random effects were applied to a random intercept and random slope on time.
ICD-9, International Classification of Diseases, Ninth-Revision; ICD-10, International Classification of Diseases, Tenth Revision.

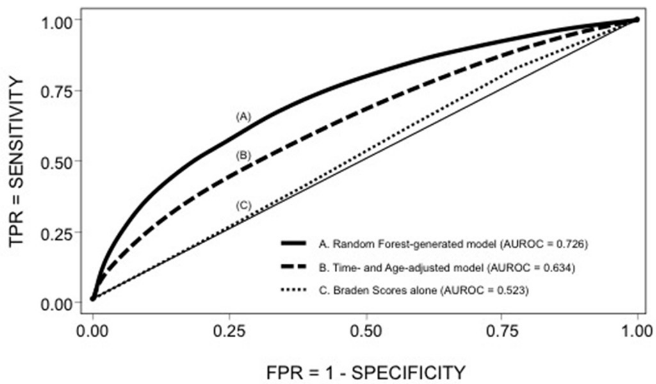


**Figure 1** Receiver operating characteristics curve (ROC) for prediction of pressure injury risk. The ROC curves and statistics for area under the ROC (AUROC) are shown for predictive algorithms and standard care after five risk assessments: (A) random forest-generated machine learning model; (B) time-adjusted and age-adjusted model; (C) standard care with the Braden scale. The sensitivity and specificity performance are greatest using the machine learning generated model. FPR, false positive rate; TPR, true positive rate.

on Patient Outcomes and Stakeholder Engagement' (PURPOSE) Advisory Council at University of Southern California, which consists of patient advocates and provider stakeholders. On completion of external peer-review of this study, results were shared with the PURPOSE Advisory Council.

## RESULTS

### Patient population

We mined a cohort of patients to develop a predictive algorithm of pressure injury risk from clinical data warehouses at four hospitals affiliated with US academic medical centres. This provided us with 94 745 patient EHRs collected over a 5-year period. We analysed 11 different combinations of sequential risk scores over time taken from a sample of 21 388 hospitalised patients who had at least 5 hospital-days of data. These patients experienced 35 001 separate hospitalisations and 169 709 transfers between units, meaning that over 63% of hospitalisations were observed in individuals with more than one admission (table 1). These EHRs also recorded 1549 individual pressure injury cases, some of which occurred to the same patient on different parts of the body or during separate admissions.

The median age of these individuals was 58 years (mean: 56.6; range: 18–108). While the cut-off for minimum length of a hospital stay was 5 days, the average hospital length of stay was 18.7 days. Patients had pressure injury risk scores collected every 12 hours using the Braden Scale, meaning that over a 5-day period there were 10 risk scores collected.

### Characterisation and optimisation of the predictive algorithm

The predictive algorithm was developed using risk scores across multiple time increments and evaluating the improved fit and predictive validity of the model with the addition of data from more time points, as well as additional predictors obtained through machine learning feature selection methods. We rejected the fixed-effects model in favour of a multilevel regression model with mixed-effects based on reduced error according to a $\chi^2$ test of the (LLRs; $p<0.001$). The multilevel model provided different estimates of coefficients in predicting the trajectory of risk scores over time, but these coefficients were preferred since the random-intercept offered subject-specific estimates of risk.

A list of 40 predictors returned from machine learning was then narrowed based on clinical meaningfulness (ie, a reasonable causal relationship to pressure injury outcomes) and fit to the model using a stepwise approach. The approach returned 20 statistically significant and clinically meaningful predictors in addition to time and age from 4 categories: (a) prescription drugs—beta blockers, electrolytes, phosphate replacement, zinc replacement, erythropoietin stimulating agents, thiazide/diuretics, analgesics/narcotics, vasopressors; (B) diagnoses—dyspnoea, hyperlipidaemia, bed confinement, spinal cord injury and chronic osteomyelitis; (c) lab orders—urinalysis, lipid panel, prealbumin, stool culture, quantitative culture and stain with anaerobes and vancomycin random assay and (d) medical orders—upper extremity venous duplex exam.

A best-fit model was determined to possess risk scores between 5 and 6 time points, whereby enough risk scores reduced error and improved fit. A plot of the ROC for both models, combined with resulting area under the ROC (AUROC) showed that 5 and 6 time points were approximately equal (AUROC=0.73). Ultimately, the marginal benefit of an additional 12-hour shift may not have outweighed the associated costs of skilled labour and other resources used for an additional risk score recording 12 hours after the fifth risk score.

### Predictive algorithm characteristics

For five risk assessments at equally timed intervals, we performed the multilevel logistic regression on the Braden Scale alone, then with time and age adjustments (table 2), and finally with the full model including random-forest-generated predictive features (table 3). A test of the LLRs for these models favoured the model with predictors generated from machine learning.

A plot of the ROC characteristics also favoured the machine learning algorithm (figure 1). In particular, at a specificity of 0.5, the machine learning model exceeded a sensitivity of 0.75; whereas the time-adjusted and age-adjusted model only reached a sensitivity of 0.596 at equal specificity. Standard care with the Braden Scale alone was lowest performing of these three information sets.

The multilevel regression approach assumed linearity in subject-specific risk trajectories. We tested whether there were elements of non-linearity in the subject-specific outcomes by applying quadratic random-effects to samples to discern whether non-linearity could

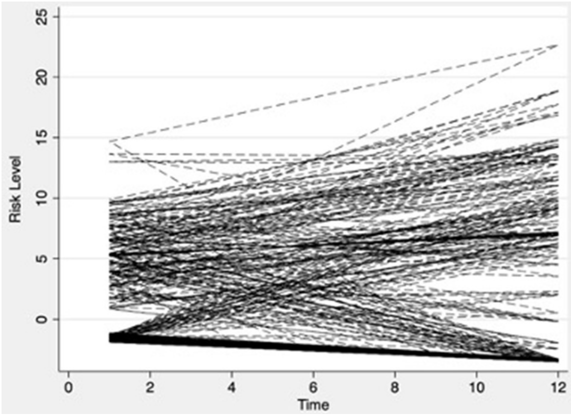


**Figure 2** Risk score trajectories for a multilevel logistic regression model with quadratic random effects. The absence of pronounced curvature in their trajectories indicates that most risk score trajectories possess linearity.

be observed. The absence of non-linearity ruled out the need for more sophisticated (ie, deep learning) approaches to statistical prediction (figure 2). The use of a generalised linear model with maximum likelihood estimates was sufficient to capture the majority of the variance in our sample and is more easily interpretable and generalisable for policy and clinical decision-makers.

### Cross-validation

Data from the hospital EHR were divided into training and testing datasets in order to conduct fivefold cross-validation on the accuracy and generalisability of the predictive algorithm. The training datasets contained 60% of the encounters, and testing dataset contained 40% of the encounters from the original sample. The test model consistently predicted the observed outcomes in the training model between 76.96% and 77.34% of the time for each of the five runs.

### Financial impact

The predictive algorithm offers improved economic efficiency, such that a hospital could recoup substantial savings on a weekly basis. Since a risk assessment can take anywhere from 5 to 15 min per patient, this could represent up to 250 labour-hours in a single 500-bed facility per day, and between 30 000 and 90 000 labour-hours per year. An average 500-bed hospital would spend up to US$99.44 per patient per day on follow-up preventive tasks. For a hospital at 100% vol capacity, this represents a weekly investment of about US$348 000 or US$18 million annually. By comparison, a predictive algorithm that limits follow-up after the fifth time point reduces costs by over 48%; reduced follow-up after the sixth time point still can save as much as 42%. The 6% gain in economic efficiency between the fifth and sixth time point could represent a savings of about US$975 000 annually without substantial losses in sensitivity or specificity.

## DISCUSSION

Using measurements commonly collected in EHRs to track risk for pressure injury, we developed an individual-specific algorithm to predict future risk and direct labour-intensive patient care. The predictive algorithm offers the potential for health systems to address the needs of patients at risk for pressure injury using modern technology. Pressure injuries remain one of the most fatal, high-cost events in healthcare despite an extent of evidence that contributes to our understanding of their effective prevention. A predictive algorithm that can direct care efficiently, and with greater sensitivity, specificity and AUROC than standard risk measures and subjective clinical judgement may increase the likelihood that providers will follow through on prevention guidelines when prescribed. Health systems should consider using predictive algorithms that introduce significant gains in AUROC, which represent a balance in gain of sensitivity and specificity, over the Braden Scale alone.

This predictive algorithm offers health systems an opportunity to cut costs on wasteful spending by over 40%, focusing efforts on patients who have higher-probability trajectories towards pressure injury incidence. The data used to train the predictive algorithm illustrate that high-risk patients are greater than 10 times as likely to develop a pressure injury compared with lower-risk patients.[29] Thus, the minority of hospitalised patients who remain immobilised, malnourished and have recurring incontinence issues will either sustain high-risk scores or establish a trajectory towards high risk.[31]

This predictive algorithm supports an ongoing mission among healthcare providers to be more prescriptive with patients when it comes to directing evidence-based care. This effort reflects the combination of descriptive patient characteristics (eg, clinical and demographic information) and predictive knowledge gained from calibration data in order to prescribe evidence-based care to individuals among a patient population most likely to benefit.[10] Prescriptive care among providers in healthcare often loses sensitivity since the prior innovations lacked accuracy to complement descriptive characteristics with predictive knowledge.[19] By improving the sensitivity of prescriptive care, providers may be more inclined to follow through on evidence-based guidelines that are time-consuming or labour intensive by knowing that this care holds good value when directed to patients who are most likely to benefit.

That being said, pressure injury prevention and incorporation of this specific predictive algorithm with an existing EHR may be the best place to start if health systems are uncertain of how to prioritise prevention of different sentinel events. Pressure injury risk is generally composed of issues related to mobility, nutrition and moisture management (ie, incontinence). The Braden Scale already adjusts for these risk factors, depending on the expertise of trained clinicians, and overlaps with risk factors for most other iatrogenic injuries for which CMS continues to withhold payments or penalise health systems.[32] For instance, mobility management is critical to preventing fall injury, deep vein thrombosis and ventilator-acquired pneumonia; nutrition is a critical component of infection mitigation; and moisture management is important when weighing the risk of infections caused by urinary catheters, *Clostridium difficile* or sepsis.[31] Health systems uncertain about where to begin a technology-driven quality improvement programme with predictive analytics could start with this tool for pressure injury prevention as a way to stabilise the risk factors for many never events, and then expand over time to manage the rates of other outcomes that continue to challenge hospital performance measures. That being said, health systems should be prepared to adjust their culture to the use of these tools—overcoming barriers such as alarm fatigue brought on by risk notifications, and common data models to perform predictive analytics using the longitudinal data structure required by this predictive algorithm are critical to successful implementation.

## Limitations

There are several limitations to this study. First, there are inherent considerations about the reliability of Braden scores using EHR data. This predictive algorithm should be used in tandem with the Braden Scale and clinical judgement to identify patients who need immediate intervention in the first several days before transitioning individual patients to follow-up triggered by the EHR.[6] Second, the retrospective nature of accessing these clinical data makes it difficult to verify the accuracy of Braden scores and other clinical data with respect to what was the true condition of the patient being observed. Third, the generalisability of this model outside of the observed academic medical centres is unknown and needs to be externally validated. Fourth, the study uses codes entered into a patient EHR in real time. There is potential bias in the entry of latter codes based on earlier diagnoses, prescriptions and orders. Fifth, given known biases of the Braden Scale towards Caucasian skin tones, there may be inherent bias in this predictive algorithm based on its calibration data, unless future studies of this algorithm can control for race and ethnicity.

Sixth, many other predictive features may exist, both within the EHR and unobserved. Our data request was limited to the listed classifications of covariates to simplify a complex data management process by as much as possible. Additionally, unobserved factors in the dataset such as race/ethnicity and sex may have also improved the model, but our IRB protocol limited access to multiple identifying factors at once.

Seventh, the choice to use logistic mixed regression for predictive analytics is a simple, supervised approach that fits within the ISPOR guideline for machine learning in outcomes research.[19] While it depends on features provided from more sophisticated selection approaches (eg, random forests), the prediction method itself is not considered a modern, advanced machine learning approach.[33] Combining coefficients derived from this study with unsupervised methods (eg, deep learning) could represent an evolution in future research; however, it does violate the transparency and interpretability criteria we set with clinical collaborators for this study.

Eighth, the NPIAP Standardised Pressure Injury Prevention Protocol Checklist recommends the use of validated tools to measure nutrition and mobility levels.[34] While such tools were not captured in the EHR used in this study, our study team and PURPOSE Advisory Council reviewed resulting predictors from the random forest that could be helpful proxies for nutrition (eg, lab orders for prealbumin and lipid panels) and mobility (eg, bed confinement and spinal cord injury). Future research could improve on this analysis by including validated tools for mobility and nutrition in predictive analytics.

Lastly, a number of unique findings in this predictive algorithm could be considered paradoxical. Stool culture orders were associated with reduced risk, perhaps because despite having risk of bacterial infection and diarrhoea, ordering this lab within 5 days of admission places patients on a correct path to reduced pressure injury risk. Prescribed electrolyte replacement was protective, whereas phosphate replacement increased risk, perhaps differentiating the underlying risk factors of patients needing one replacement compared with the other. Exceptionally high ORs for bed confinement and spinal cord injury, despite small sample sizes, are worth additional cross-validation with other data samples given the small sample sizes of these data. While there is potential that some of these sets of predictors are highly correlated, our test for autocorrelation or non-parametric fit of the underlying covariance structure of the model returned no alternative findings.

## Conclusions

Predictive analytics is now commonplace in hospitals to identify patients who are high risk for costly, harmful events and take steps to mitigate the circumstances of their risk.[35] Pressure injuries are particularly important in this light, causing 60 000 deaths per year and costing US healthcare systems over US$25 billion annually.[36] While pressure injury prevention represents a cost-effective alternative to treating the outcome, it requires investment in clinical staff and material resources to follow clinical practice guidelines. A predictive algorithm such as this could limit direct follow-up for pressure injury prevention to those patients of greatest need after several days of data collection, thereby reducing the costs of prevention by 40%–50%. These savings for hospitals would create financial bandwidth to justify the transition to EHR-based warning systems that use predictive algorithms to trigger response for patients at high risk of iatrogenic injury. Furthermore, pressure injury prevention programmes offer substantial financial savings that can be recouped to invest in other patient safety programmes in the future.

**Author affiliations**
[1]Department of Pharmaceutical & Health Economics, University of Southern California Mann School of Pharmacy & Pharmaceutical Sciences, Los Angeles, CA, USA
[2]Stage Analytics, Suwanee, GA, USA
[3]The Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California, Los Angeles, CA, USA
[4]Department of Surgery, USC Keck School of Medicine, Los Angeles, California, USA
[5]University Hospitals of Cleveland, Shaker Heights, Ohio, USA
[6]Anesthesiology and Critical Care Medicine, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA
[7]Department of Computer Science, Johns Hopkins University Whiting School of Engineering, Baltimore, Maryland, USA
[8]Department of Health Policy & Management, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA

**Twitter** William V Padula @DrWmPadula

**Contributors** WVP, DGA, PJP and SS contributed to this manuscript in accordance with ICMJE: WVP, DGA, PJP and SS were involved in the concept and design of the study, critical revision of the paper for important and intellectual content, supervision and administrative support of the study. WVP was involved in the acquisition of data, statistical analysis and interpretation, and drafting of the manuscript. WVP and DGA were involved in obtaining funding for the resulting study. WVP is responsible for the overall content as the guarantor of the work, and accepts full responsibility for the work and/or conduct of the study, had access to the data, and controlled the decision to publish.

**Funding** This study was supported by a grant from the National Institutes of Health Office of Extramural Research (KL2 TR001854, PI: WVP).

**Competing interests** WVP declares personal fees and equity holdings from Stage Analytics. SS declares equity holdings from Bayesian Health.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Ethics approval** Ethics approval was obtained from University of Chicago Medicine protocol #IRB14-0869; University of Southern California protocol #UP-20-00140.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request. All data, code and materials using in this analysis are available freely on request to WVP.



**ORCID iDs**
William V Padula http://orcid.org/0000-0003-1161-6954
David G Armstrong http://orcid.org/0000-0003-1887-9175
Peter J Pronovost http://orcid.org/0000-0002-9740-3775
Suchi Saria http://orcid.org/0000-0002-7667-5210

## REFERENCES

1 AHRQ National Scorecard on Hospital-Acquired Conditions. Updated baseline rates and preliminary results 2014–2017 secondary updated baseline rates and preliminary results 2014–2017. 2019. Available: https://www.ahrq.gov/hai/pfp/index.html
2 Reddy M, Gill SS, Kalkar SR, *et al*. Treatment of pressure ulcers: a systematic review. *JAMA* 2008;300:2647–62.
3 Lyder CH, Ayello EA. Pressure ulcers: A patient safety issue. In: Hughes RG, ed. *Patient Safety and Quality: An Evidence-based Handbook for Nurses*. Rockville, MD: Agency for Healthcare Research and Quality, 2008.
4 Padula WV, Delarmente BA. The National cost of hospital-acquired pressure injuries in the United States. *Int Wound J* 2019;16:634–40.
5 Nussbaum SR, Carter MJ, Fife CE, *et al*. An economic evaluation of the impact, cost, and Medicare policy implications of chronic Nonhealing wounds. *Value in Health* 2018;21:27–32.
6 Qaseem A, Mir TP, Starkey M, *et al*. Risk assessment and prevention of pressure ulcers: a clinical practice guideline from the American college of physicians. *Ann Intern Med* 2015;162:359–69.
7 Creditor MC. Hazards of hospitalization of the elderly. *Ann Intern Med* 1993;118:219–23.
8 Haesler E, ed. *European Pressure Ulcer Advisory Panel, National Pressure Ulcer Advisory Panel, Pan-Pacific Pressure Injury Alliance. Prevention and Treatment of Pressure Ulcers/Injuries: Quick Reference Guide*. Boston, MA: EPUAP/NPIAP/PPPIA, 2019.
9 Braden BJ, Bergstrom N. Predictive validity of Braden scale for pressure sore risk in a nursing home population. *Res Nurs Health* 1994;17:459–70.
10 Padula WV, Gibbons RD, Pronovost PJ, *et al*. Using clinical data to predict high-cost performance coding issues associated with pressure ulcers: a Multilevel cohort model. *J Am Med Inform Assoc* 2017;24:e95–102.
11 Rajkomar A, Oren E, Chen K, *et al*. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018;1:18.
12 Churpek MM, Yuen TC, Winslow C, *et al*. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit Care Med* 2016;44:368–74.
13 Ghanem-Zoubi NO, Vardi M, Laor A, *et al*. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. *Crit Care* 2011;15:R95.
14 Barriere SL, Lowry SF. An overview of mortality risk prediction in sepsis. *Crit Care Med* 1995;23:376–93.
15 Jones AE, Trzeciak S, Kline JA. The sequential organ failure assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency Department presentation. *Crit Care Med* 2009;37:1649–54.
16 Vorwerk C, Loryman B, Coats TJ, *et al*. Prediction of mortality in adult emergency Department patients with sepsis. *Emerg Med J* 2009;26:254–8.
17 Henry KE, Hager DN, Pronovost PJ, *et al*. A targeted real-time early warning score (Trewscore) for septic shock. *Sci Transl Med* 2015;7:299ra122.
18 Alderden J, Pepper GA, Wilson A, *et al*. Predicting pressure injury in critical care patients: A machine-learning model. American Journal of critical care: an official publication. *American Association of Critical-Care Nurses* 2018;27:461–8.
19 Padula WV, Kreif N, Vanness DJ, *et al*. Machine learning methods in health economics and outcomes research-the PALISADE checklist: a good practices report of an ISPOR task force. *Value Health* 2022;25:1063–80.
20 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594.
21 Hedeker D, Gibbons RD. *Longitudinal Data Analysis*. Hoboken, NJ: Jon Wiley & Sons, Inc, 2006.
22 Padula WV, Blackshaw L, Brindle CT, *et al*. An approach to acquiring, Normalizing, and managing EHR data from a clinical data repository for studying pressure ulcer outcomes. *J Wound Ostomy Continence Nurs* 2016;43:39–45.
23 Agency for Healthcare Research and Quality (AHRQ). Patient safety indicator 03 (PSI 03) pressure ulcer rate. AHRQ quality indicators (AHRQ QI) ICD-10-CM/PCS specification 2021 March. 2020.
24 Edsberg LE, Black JM, Goldberg M, *et al*. Revised national pressure ulcer advisory panel pressure injury staging system: revised pressure injury staging system. *J Wound Ostomy Continence Nurs* 2016;43:585–97.
25 Le NQK, Do DT, Nguyen T-T-D, *et al*. A sequence-based prediction of Kruppel-like factors proteins using Xgboost and Optimized features. *Gene* 2021;787.
26 James G, Witten D, Hastie T, *et al*. An introduction to statistical learning. In: *An Introduction to Statistical Learning*. New York, NY: Springer, 2013.
27 Nguyen HS, Ho DKN, Nguyen NN, *et al*. Predicting EGFR Mutation status in non-small cell lung cancer using artificial intelligence: A systematic review and meta-analysis. *Acad Radiol* 2024;31:660–83.
28 Allman RM. Pressure ulcers in the elderly. *N Engl J Med* 1989;320:850–3.
29 Padula WV, Pronovost PJ, Makic MBF, *et al*. Value of hospital resources for effective pressure injury prevention: a cost-effectiveness analysis. *BMJ Qual Saf* 2019;28:132–41.
30 Padula WV, Mishra MK, Makic MBF, *et al*. Improving the quality of pressure ulcer care with prevention: a cost-effectiveness analysis. *Med Care* 2011;49:385–92.
31 Padula WV, Armstrong DG, Goldman DP. Complexity bias in the prevention of iatrogenic injury: why specific harms may inhibit performance. *Mayo Clin Proc* 2022;97:221–4.

32 Padula WV, Black JM, Davidson PM, *et al*. Adverse effects of the Medicare PSI-90 hospital penalty system on revenue-neutral hospital-acquired conditions. *J Patient Saf* 2020;16:e97–102.
33 Lim T-S, Loh W-Y, Shih Y-S. A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification Algorithms. *Mach Learn* 2000;40:203–28.
34 Pittman J, Black JM, de Jesus A, *et al*. The standardized pressure injury prevention protocol checklist 2.0: content validation. *J Adv Nurs* November 23, 2023.
35 Patel V, Jamoom E, Hsiao CJ, *et al*. Variation in electronic health record adoption and readiness for meaningful use: 2008-2011. *J Gen Intern Med* 2013;28:957–64.
36 Padula WV, Pronovost PJ. Addressing the Multisectoral impact of pressure injuries in the USA, UK and abroad. *BMJ Qual Saf* 2019;28:132–41.