Received: 6 November 2021 | Revised: 17 January 2022 | Accepted: 18 January 2022

DOI: 10.1111/iwj.13764

ORIGINAL ARTICLE

IWJ WILEY

Development and validation of a machine learning algorithm–based risk prediction model of pressure injury in the intensive care unit

Jie Xu[1] | **Danxiang Chen**[2] | **Xiaofang Deng**[3] | **Xiaoyun Pan**[1] | **Yu Chen**[4] | **Xiaoming Zhuang**[1] | **Caixia Sun**[4]

[1]Department of Thoracic Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

[2]Department of Breast Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

[3]Nursing department, Wenzhou Medical University, Wenzhou, China

[4]Nursing Department, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

Correspondence
Caixia Sun, Nursing Department, The First Affiliated Hospital of Wenzhou Medical University, Nanbaixiang, Wenzhou, Zhejiang, 325000, China.
Email: scx_nursing@163.com

Funding information
Natural Science Foundation of Zhejiang Province, Grant/Award Number: LQ17H260004; Science and Technology Association of Wenzhou, Grant/Award Number: rxk02; Science and Technology Foundation of Wenzhou, Grant/Award Number: Y20180038

Abstract

The study aimed to establish a machine learning–based scoring nomogram for early recognition of likely pressure injuries in an intensive care unit (ICU) using large-scale clinical data. A retrospective cohort study design was employed to develop and validate a top-performing clinical feature panel accessibly in the electronic medical records (EMRs), which was in the mode of a quantifiable nomogram. Clinical factors regarding demographics, admission cause, clinical laboratory index, medical history and nursing scales were extracted as risk candidates. The performance improvement was based on the application of the machine learning technique, comprising logistic regression, decision tree and random forest algorithm with five-fold cross-validation (CV) technique. The comprehensive assessment of sensitivity, specificity and the area under the receiver operating characteristic curve (AUROC) was considered in the evaluation of predictive performance. The receiver operating characteristic curves revealed the top performance for the logistic regression model in respect to machine learning improvement, achieving the highest sensitivity and AUC among three types of classifiers. Compared against the 23-point Braden scale routinely recorded online, an incorporated nomogram of logistic regression model and Braden scale achieved the best performance with an AUC of 0.87 ± 0.07 and 0.84 ± 0.05 in training and test cohort, respectively. Our findings suggest that the machine learning technique potentiated the limited predictive validity of routinely recorded clinical data on pressure injury development during ICU hospitalisation. Easily accessible electronic records held the potentials to substitute the traditional Braden score in the prediction of pressure injury in intensive care unit. Preoperative prediction of pressure injury facilitates the exemption from the severe consequences.

Abbreviations: AUROC, area under the receiver operating characteristic curve; CV, cross validation; DCA, decision curve analysis; EMR, electronic medical record; ICU, intensive care unit; PI, pressure injury.

KEYWORDS
electronic medical record, machine learning algorithm, pressure injury, risk prediction

Key Messages

- easily accessible electronic records held the potentials to substitute the traditional Braden score in the prediction of pressure injury in intensive care unit
- preoperative prediction of pressure injury facilitates the exemption from the severe consequences

1 | INTRODUCTION

Pressure injury (PI), defined as the local damage to the skin and underlying soft tissue owing to the prolonged pressure, still hangs in doubt in the medical field.[1] PIs were reported to frequently occur in the intensive care unit (ICU) department, where patients undertook a heavy burden in both expenditure and life.[2,3] Considering the burdensome consequences of PI, particularly in ICU, it is compelling to identify the most relevant risk factors and perform etiological prevention.[4,5]

An upsurge of PI incidence in aggressive ICU was observed in recent years.[6] Owing to the abysmal prognosis and curative rates, there was a diffuse thought to take a holistic preventive approach, particularly in the nursing field.[7,8] The Braden scale to predict and assess the hazards developing PIs is worldwide used in nowadays clinical settings, which was attributed to a kind of bedside nursing.[9] This scale constituted of six subscales, including activity-, friction-, mobility-, moisture-, nutrition- and sensory perception-related scoring.[10] However, the Braden scale demonstrated an insufficient performance and labour-consuming trait, so that numerous studies devoted to finding better-performing estimates to require less from nursing resources.[11,12] Since most machine learning algorithms were not presented intuitively, previous studies held predilection for logistic regression analysis, which quantified the risk formulation for clinicians to easily evaluate the risk probabilities.[13,14]

In this study, we generated fresh insight into the fusion of machine learning and risk prediction to develop a brand-new model for predicting the risk probability of PIs in ICU.

2 | METHODS

2.1 | Aim and study design

Aiming at a specialised risk identification of the PI development in the ICU population, we conducted a retrospective observational cohort based on the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines,[15] using the existing electronic medical records for data mining, which took place in the ICU department with a 60-bed allocation. To that end, the study generated various machine learning–based clinical models and sequentially anatomise the clinical value by comparison between the novel clinical model and the traditional Braden scale. This study was approved by the ethics committee of the First Affiliated Hospital of Wenzhou (Approval No. 2020052) in China.

2.2 | Participant recruitment and incident tracking

A total of 618 samples was gathered upon admission to ICU during their hospital stay from 1 January 2020, through 30 June 2021. The inclusion criteria for the study permitted all patients in ICU, free of the limitation of age and gender. Amongst the patients in ICU, those holding the purpose of perioperative observation within the 24 hours of post operation, and palliative care with little chance of being cured, were excluded for equilibrium. Subsequently, the whole electric medical records were searched to track PI incidents occurred in ICU. And patients whose discharge diagnosis of PI were assigned with the PI label (N = 206), while those with entry diagnosis on ICU admission of PI were assumed as non-PI developed in ICU (N = 412), regardless of their previous history of PI in the general ward.

According to the time interval from admission in ICU to PI occurrence, we forward demarcated the PI-developed participants into two subgroups: early-developed cohort within 24 hours in ICU (N = 53) and late-developed cohort after 24 hours in ICU (N = 153).

During the study period, complete clinical information referred to commonly adopted demographics, a head-to-toe PI nursing assessment, related clinical laboratory index and medical history was collected upon ICU admission within 24 hours. For the late-developed

cohort, a compromised indicator-recorded resolution 1 week in advance was adopted in avoidance of the interference brought by floating illness severity in a long stay of ICU encounter. The total 618 samples gathered were each assigned a unique serial number, and all their clinical information was de-identified by a third-party staff before investigators' accession to these data.

2.3 | Cohort formulation

At the time point of analysis, 70% and 30% of the samples were split into training and test cohorts by principle. In the practical application, the ratio of division in the whole samples was shared in the subgroups with opposite PI-related outcomes to separately gather the subpopulation to constitute the training and test cohorts, functioning in the equilibrium of event occurrence. The training set further conducted a quintile subdivision for parameter tuning. The procedure of five fold inner cross-validation, where one fold iteratively served as inner validation set in each round, while the remaining four folds as the inner training set were repeated five times till every fold experienced inner validation. The training and inner validation set participated in the model construction and parameter optimization, and the test cohort independent of the inner loop was simply for the performance test.

2.4 | Feature engineering and outcome

Preliminary screening of the candidate features was conducted in the range of electronic medical records and nursing records, which were literature-reportedly related to the PI outcome. Regarding the data integrity and acquirability in EHR, we finally compiled 31 clinical parameters relevant to PI development. The alternatives were outlined as followed: demographics (age, body mass index [BMI], gender), classification of the admission cause (chronic consumptive diseases, acute critical diseases), clinical laboratory index (haemoglobin, haematocrit, blood urea nitrogen, creatinine, blood lactate, bilirubin, albumin), medical history (history of hypertension, diabetes mellitus, stroke, disturbance of consciousness, peripheral vascular disease, mechanical ventilation dependency, hypnotics and sedatives, analgesic, vasoactive agent, surgical history) and nursing scale (Braden score: moisture subscore, sensitivity perception subscore, activity subscore, mobility subscore, nutrition subscore, friction subscore and Glasgow Coma Scale).

Those dynamically changing parameters along with the ICU hospitalisation were arranged chronologically. Once PI was developed, the nearest point of records before PI development was input as vital metrics for analysis. For those negative control groups, data collection was conducted within 24 hours on the admission of ICU. Data pre-processing was further conducted with the advent of missing values. Since the machine learning algorithms properly handles the missing values, we just keep missing data as a valid value in the decision tree (DT) and RF algorithm. The intelligentialised algorithm would take the missing value as an independent attribute. In the logistic regression modelling, imputation was a better choice premised on the missing at random (MAR) than leaving them as is, particularly in the situation of no sense made by the missing value below the threshold. Based on the characteristics of missing variables, the average value was applied to fill up the missing value of continuous variables and modes for categorical variables. Then, a qualified dataset of clinical information matched to the outcome variables and nursing scale scores came into being.

We mainly focussed on the development of pressure injuries in the ICU, which appeared to be the high-hazard department in the hospital. Based on the continued exposure in general ward spreading and carried to ICU, we proposed a reclassification according to the ICU stay length up to PI onset. A large collection of PI events were recorded in the nursing board on EHR with their injured body part. The rest was searched according to the specialised management aiming at PI.

2.5 | Outcome-related risk factors identification and machine learning–based model construction

After the initial stage of pooling candidates with their pre-specified clinical information in a dataset, the correlation analysis embarked from the univariate analysis, to elucidate an independent panel of outcome-related risk factors in general. Given the indetermination of the causality, we lift a restriction in statistical power as much as possible to provide more assumptions for potential risk factors in the machine learning techniques.

Regarding machine learning, we performed logistic regression (LR), DT[16] and random forest (RF).[17,18] R package for implementation comprised glmnet for LR, the recursive partitioning and regression trees (rpart) for DT and randomForest for RF. Multivariate logistic regression was based on the abovementioned univariate analysis to further eliminate variables with collinearity. The intuitive visualisation of the linearity relationship between risk factors and PI outcome by the LR technique was, to some extent, at a cost in the interpretability of non-linear relationships. A delicate-modelled tree experienced pruning approaches and developed into a mature

model with minimum intra-class error and maximum post-partition purity, which was the essence of DT.[19] RF, known as an ensemble learning technique, was proposed as bootstrap aggregating of DTs.[20,21] DT and RF algorithms, previously reported to manage the missing data modelling by themselves, actually master a hard substitution rule, orchestrating a more vulnerable discordance owing to an appreciable fitting degree. To relieve the overfitting brought by the complexity particularly in the small sample, cross-validation was essential to address these imbalances of predictive validity between training and test cohorts by maximising the parameter stability. Entropy or Gini impurity of predictor variables was perceived as the impurity optimisation in a step-by-step manner, which substituted P-values or coefficients in the DT model.

Top important feature panels generated by different machine learning algorithms were simultaneously used to construct a machine learning–based prediction model for PI prediction in ICU. Logistic regression formulated the linear connection via the corresponding coefficients of the top features, which shed light on the induction process from the unimpressive features into a destructive PI outcome. DT analysis depicted a complex tree model, which constituted cascaded decision nodes and leaf nodes split from root nodes. Root nodes, including different split nodes, were recursively tested upward and downward to reach a maximum depth of the tree model, decided by the probability of incorrect classification when randomly selected. RF visualised a full view of various DT models and averaged the output of response variables accumulated by a single DT.

2.6 | Nomogram development and validation

The selection criteria toward prediction performance and model discrimination were are determined with the metrics, including AUROC, sensitivity, specificity and so on. In consideration of the performance orientation, we tended to choose a lower threshold of specificity to relieve the terrible trauma once PI events developed. We designated the putative effect yielding the highest sensitivity and a relative lower specificity above 0.8 and considered a combination between the best-performing machine learning-based prediction model and the current Braden scale predisposition to PI. Owing to the prespecified threshold on performance metrics, logistic regression was modelled together with the Branden scale to build a scoring nomogram. And then inter-comparison between the three machine learning–based models, the traditional Braden model and the novel nomogram were conducted in predictive performance, calibration and discrimination.

The standard evaluation process began with the metrics, named sensitivity, specificity, accuracy, precision, positive predictive value (PPV), negative predictive value (NPV) and area under the curve (AUC). Besides, accuracy was reflected in the comparison of the index integrated discrimination improvement[22] and the net reclassification improvement (NRI). In terms of calibration capacity, calibration curves were plotted to explore

TABLE 1 Summary of the included potential risk factors

Categories	Specific item
Classification of the admission cause	Chronic consumptive diseases, acute critical diseases
Clinical laboratory index	Haemoglobin, haematocrit, blood urea nitrogen, creatinine, blood lactate, bilirubin, albumin
Demographics	Age, BMI, gender
Medical history	History of hypertension, diabetes mellitus, stroke, disturbance of consciousness, peripheral vascular disease, mechanical ventilation dependency, hypnotics and sedatives, analgesic, vasoactive agent, surgical history
Nursing scale	Braden score (moisture subscore, sensitivity perception subscore, activity subscore, mobility subscore, nutrition subscore, friction subscore), Glasgow Coma Scale

TABLE 3 Key features selected by different selection methods

Method	Number of features	Entries
Univariate/ multivariate logistic regression	7	Age, history of diabetes mellitus, history of stroke, peripheral vascular disease, mechanical ventilation, GCS, surgical history.
Decision tree	4	Bilirubin, peripheral vascular disease, GCS, surgical history
Random forest	6	Bilirubin, age, history of diabetes mellitus, peripheral vascular disease, GCS, surgical history

Abbreviation: GCS, Glasgow Coma Scale.

TABLE 2 Baseline characteristics of training and validation cohort and univariate analysis of risk factors

Features	Training (N = 472)			Test (N = 146)		
	PI developed (N = 159)	Not developed (N = 313)	*P*	PI developed (N = 47)	Not developed (N = 99)	*P*
Age, mean ± SD	64.5 ± 8.7	56.4 ± 7.4	<.01	63.7 ± 5.8	56.9 ± 6.1	<.01
Sex (%)			.037			.028
Female	96 (60.4)	147 (47.0)		29 (62.4)	45 (45.9)	
Male	63 (39.6)	166 (53.0)		18 (37.6)	54 (54.1)	
BMI, no.(%)			<.001			<.001
<18	86 (54.1)	89 (28.4)		24 (51.4)	27 (27.5)	
<24	46 (28.9)	154 (49.2)		12 (25.9)	47 (47.3)	
≥24	27 (17.0)	70 (22.4)		11 (22.7)	25 (25.2)	
History of diabetes mellitus, no. (%)			<.01			<.01
Yes	56 (35.2)	67 (21.4)		16 (33.9)	19 (19.5)	
No	103 (64.8)	246 (78.6)		31 (66.1)	80 (80.5)	
History of hypertension, no. (%)			.015			.026
Yes	115 (72.3)	210 (67.1)		34 (71.8)	69 (69.4)	
No	44 (27.7)	113 (36.1)		13 (28.2)	30 (30.6)	
History of stroke, no. (%)			<.001			<.001
Yes	37 (23.3)	22 (7.0)		10 (22.3)		
No	122 (76.7)	291 (93.0)		37 (77.7)	93 (94.1)	
Peripheral vascular disease, no. (%)			<.001			<.001
Yes	87 (54.7)	97 (31.0)		25 (53.5)	33 (33.4)	
No	72 (45.3)	216 (69.0)			66 (66.6)	
Disturbance of consciousness, no. (%)			<.001			<.001
Yes	69 (43.4)	61 (19.5)		21 (45.1)	22 (22.5)	
No	90 (56.6)	252 (80.5)		26 (54.9)	77 (77.5)	
Mechanical ventilation, no. (%)			<.001			<.001
Yes	92 (57.9)	97 (31.0)		26 (56.3)	33 (33.5)	
No	67 (42.1)	216 (69.0)		21 (43.7)	66 (66.5)	
History of hypnotics and sedatives, no. (%)			.045			.038
Yes	31 (19.5)	38 (12.1)		7 (14.9)	12 (12.3)	
No	128 (80.5)	275 (87.9)		40 (85.1)	87 (87.7)	

(Continues)

TABLE 2 (Continued)

Features	Training (N = 472)			Test (N = 146)		
	PI developed (N = 159)	Not developed (N = 313)	*P*	PI developed (N = 47)	Not developed (N = 99)	*P*
History of analgesic, no. (%)			.032			.043
Yes	74 (46.5)	127 (40.6)		21 (44.8)	41 (41.3)	
No	85 (53.5)	186 (59.4)		26 (55.2)	58 (58.7)	
History of vasoactive agent, no. (%)			<.001			<.001
Yes	86 (54.1)	103 (32.9)		25 (53.8)	34 (34.2)	
No	73 (45.9)	230 (67.1)		22 (46.2)	65 (65.8)	
Hg, median (IQR), (g/L)	94 (80, 112)	93 (79, 115)	.788	96 (84, 115)	94 (81, 109)	.589
Haematocrit, mean ± SD, (%)	33.6 ± 5.4	31.4 ± 4.7	.045	33.9 ± 4.2	31.1 ± 3.9	.039
PLT, median (IQR), (×10^9/L)	147 (125, 157)	128 (117, 136)	<.001	145 (132, 156)	126 (114, 133)	<.001
D-Dimer, median(IQR), (mg/L)	0.8 (0.5, 1.2)	0.5 (0.3, 0.9)	<0.01	0.7 (0.5, 1.1)	0.4 (0.3, 0.7)	<.01
BUN, median(IQR), (mmol/L)	9 (6, 14)	7 (5, 12)	.001	10 (7, 13)	7 (5, 10)	.001
Creatinine, median (IQR), (umol/L)	95 (64, 161)	81(55, 160)	.327	93 (68, 154)	84(64, 139)	.486
Blood lactate, median (IQR), (mmol/L)	3 (2.1,3.6)	1.6 (1.3,2.5)	<.001	3 (1.9,3.8)	1.8 (1.2,2.3)	<.001
Bilirubin, mean ± SD, (umol/L)	35 (30.8, 42.5)	21 (18.1, 23.6)	<.001	33 (29.5, 38.5)	20 (16.5, 31.4)	<.001
Albumin, mean ± SD, (g/L)	32.6 ± 5.8	31.9 ± 6.9	.437	33.4 ± 3.8	31.7 ± 2.9	.326
GCS, mean ± SD	8.4 ± 3.1	14.4 ± 2.4	<.001	8.2 ± 2.7	13.4 ± 2.6	<.001
Braden score, mean ± SD	12.5 ± 5.1	19.5 ± 1.3	<.001	12.7 ± 5.0	19.1 ± 1.7	<.001
Moisture subscore, mean ± SD	2.4 ± 1.1	3.1 ± 0.9	<.001	2.5 ± 1.4	3.1 ± 1.1	<.001
Sensitivity perception subscore, mean ± SD	2.1 ± 0.7	3.5 ± 1.1	<.001	2.3 ± 0.9	3.4 ± 1.1	<.001
Activity subscore, mean ± SD	1.6 ± 0.8	3.4 ± 1.5	<.001	1.4 ± 0.7	3.1 ± 1.3	<.001
Mobility subscore, mean ± SD	2.2 ± 0.4	2.9 ± 1.1	<.001	2.1 ± 0.6	2.6 ± 1.3	<.001
Nutrition subscore, mean ± SD	1.9 ± 1.1	3.3 ± 0.7	<.001	1.7 ± 1.3	3.2 ± 0.8	<.001
Friction subscore, mean ± SD	1.5 ± 0.6	2.0 ± 0.6	<.001	1.4 ± 0.7	1.9 ± 0.8	<.001
Classification of the admission cause, no. (%)			.048			.043
Chronic consumptive diseases	119 (74.8)	174 (55.6)		34 (72.7)	53 (53.6)	
Acute critical diseases	40 (25.2)	139 (44.4)		13 (27.3)	46 (46.4)	
Surgical history, no. (%)			.034			.041
Yes	124 (78.0)	181 (57.8)		36 (77.3)	58 (58.4)	
No	35 (22.0)	132 (42.2)		11 (22.7)	41 (41.6)	

Abbreviations: BUN, blood urea nitrogen; GCS, Glasgow Coma Scale; Hg, haemoglobin; IQR, interquartile range; PI, pressure injury; PLT, platelet; SD, standard deviation.

the relevancy to which predicted probabilities approximated actuality, where a near 45° line denoted the well-performed model. Clinical efficiency was assessed by the DCA plot, where the net benefits were embodied in a quantified area under the curve to be assessed.[23] Performance evaluations were conducted in the 'rms' and 'rmda' package of Rstudio software. The *P*-value for statistical significance was set at .05.

2.7 | Statistical Methods

Continuous variables with normal distribution were presented with mean ± SD, while categorical variables in percentage were shown as median (interquartile range). Categorical variables were compared by the chi-square or Fisher's exact tests, while continuous variables were by the Mann–Whitney *U*-test. And the threshold of significance was a two-sided *P*-value less than .05.

3 | RESULTS

3.1 | Clinicopathologic characteristics

A total of 618 patients in ICU were enrolled in the analysis, 206 patients of whom with PIs were cases and 412 patients were controls. Among the PI cases, 53 patients developed PIs early within 24 hours in ICU, and the rest 153 patients developed after 24 hours since ICU admission. We gathered five panels of prediction parameters based on the recorded information online, and literature investigation, comprising of demographics, classification of the admission cause, clinical laboratory index, medical history and nursing scale. Details of variable panels are presented in Table 1.

3.2 | Univariate analysis of risk factors

Univariate analyses for screening out a panoply of risk factors roughly showed that creatinine, bilirubin and haemoglobin were not statistically significant with the occurrence of PIs in ICU (all $P > .05$; Table 2). Thus, we eliminated the above factors, and the rest features were entered into further analysis.

3.3 | Prediction performance of machine learning–based models

The basic model was constructed based on the admissive features by the univariate analysis using a different machine learning algorithm. In general, seven (multivariate regression), four DT and six RF risk factors were separately incorporated to develop the prediction models after different algorithms (Table 3). Peripheral vascular disease, GCS and surgical history were identified as significant PI predictors shared by all three basic models. Age, bilirubin and history of diabetes mellitus were perceived as significant in the two models. As shown in Table 4, the powerful predictive capacity was evidence by AUROC in three basic models (training cohort: LR: 0.82 ± 0.08; DT: 0.74 ± 0.06; RF: 0.76 ± 0.03; test cohort: LR: 0.81 ± 0.05; DT: 0.72 ± 0.04; RF: 0.77 ± 0.05.).

TABLE 4 Predictive performance of each model on pressure injury possibility

Model		Cohort	Acc, mean (SD)	Pre, mean (SD)	F1, mean (SD)	Sen, mean (SD)	Spe, mean (SD)	PPV, mean (SD)	NPV, mean (SD)	AUC, mean (SD)
Clinical	LR	Training	0.75 (0.06)	0.66 (0.05)	0.60 (0.04)	0.63 (0.06)	0.87 (0.04)	0.66 (0.03)	0.79 (0.11)	0.82 (0.08)
		Test	0.73 (0.07)	0.65 (0.08)	0.58 (0.07)	0.61 (0.04)	0.86 (0.06)	0.65 (0.05)	0.77 (0.05)	0.81 (0.05)
	DT	Training	0.65 (0.05)	0.63 (0.09)	0.63 (0.10)	0.39 (0.05)	0.81 (0.05)	0.63 (0.11)	0.81 (0.06)	0.74 (0.06)
		Test	0.62 (0.07)	0.63 (0.07)	0.62 (0.07)	0.38 (0.08)	0.80 (0.06)	0.62 (0.08)	0.82 (0.07)	0.72 (0.04)
	RF	Training	0.71 (0.03)	0.61 (0.04)	0.48 (0.03)	0.55 (0.05)	0.91 (0.04)	0.61 (0.03)	0.74 (0.07)	0.76 (0.03)
		Test	0.68 (0.07)	0.62 (0.07)	0.46 (0.05)	0.57 (0.06)	0.89 (0.03)	0.59 (0.06)	0.75 (0.08)	0.77 (0.05)
Braden		Training	0.65 (0.05)	0.58 (0.07)	0.59 (0.08)	0.60 (0.07)	0.81 (0.04)	0.58 (0.05)	0.79 (0.06)	0.75 (0.07)
		Test	0.63 (0.07)	0.59 (0.11)	0.54 (0.05)	0.57 (0.08)	0.83 (0.05)	0.54 (0.07)	0.78 (0.04)	0.76 (0.04)
Nomogram (LR + Braden)		Training	0.79 (0.09)	0.66 (0.07)	0.52 (0.06)	0.61 (0.07)	0.82 (0.04)	0.68 (0.07)	0.78 (0.04)	0.91 (0.07)
		Test	0.78 (0.07)	0.62 (0.06)	0.49 (0.07)	0.59 (0.04)	0.81 (0.07)	0.65 (0.05)	0.77 (0.08)	0.88 (0.05)

Abbreviations: DT, decision tree; LR, logistic regression; RF, random forest.

TABLE 5 Multivariate logistic analysis of risk factors for pressure injury in intensive care unit (ICU), which were of statistical significance in the univariate analysis

					Early to develop PI (N = 53)		Late to develop PI (N = 153)	
Feature	**Coefficient**	**SD**	**OR[a]**	***P***	**OR[b]**	***P***	**OR[c]**	***P***
Age	1.28	0.231	1.44 (1.31-1.49)	<.01	1.41 (1.38-1.46)	<.01	1.45 (1.32-1.56)	<.01
BMI	-	-	0.72 (0.66-0.82)	.058	0.73 (0.66-0.85)	.028	0.71 (0.63-0.84)	.076
History of diabetes mellitus	0.75	0.318	1.42 (1.28-1.59)	<.01	1.41 (1.36-1.49)	<.001	1.44 (1.27-1.63)	<.01
History of hypertension	-	-	1.13 (0.97-1.21)	.342	1.04 (0.92-1.10)	.237	1.19 (0.93-1.21)	.314
History of stroke	1.44	0.187	1.48 (1.36-1.59)	<.001	1.47 (1.33-1.57)	<.001	1.48 (1.37-1.61)	<.001
peripheral vascular disease	2.36	0.204	1.75 (1.71-1.94)	<.001	1.73 (1.68-1.77)	<.001	1.75 (1.72-1.97)	<.001
Disturbance of consciousness	-	-	1.26 (1.17-1.36)	.591	1.22 (1.16-1.34)	.439	1.27 (1.22-1.38)	.335
mechanical ventilation	0.32	0.431	1.17 (1.07-1.24)	.043	1.26 (1.19-1.31)	.014	1.13 (0.97-1.10)	.212
History of hypnotics and sedatives	-	-	1.22 (1.10-1.26)	.628	1.18 (1.13-1.27)	.725	1.23 (1.15-1.32)	.673
History of analgesic	-	-	1.12 (1.06-1.17)	.624	1.07 (1.03-1.15)	.547	1.16 (1.07-1.17)	.632
History of vasoactive agent	-	-	1.08 (1.02-1.18)	.537	1.05 (1.01-1.19)	.429	1.07 (1.02-1.23)	.574
HCT	-	-	1.35 (1.27-1.39)	.362	1.31 (1.24-1.39)	.327	1.37 (1.28-1.43)	.368
PLT	-	-	1.35 (1.29-1.39)	.147	1.43 (1.35-1.53)	.129	1.37 (1.26-1.45)	.138
D-Dimer	-	-	1.21 (1.15-1.28)	.261	1.17 (1.14-1.27)	.214	1.27 (1.17-1.31)	.271
BUN	-	-	1.35 (1.31-1.39)	.302	1.21 (1.16-1.28)	.294	1.38 (1.24-1.42)	.018
blood lactate	-	-	1.38 (1.32-1.43)	.105	1.30 (1.24-1.38)	.204	1.42 (1.35-1.51)	.174
bilirubin	-	-	1.06 (1.02-1.17)	.061	1.03 (0.99-1.10)	.117	1.15 (1.03-1.19)	.052
Albumin	-	-	0.74 (0.79-0.82)	.074	0.58 (0.49-0.67)	<.01	0.88 (0.71-1.01)	.106
GCS	−1.61	0.391	0.56 (0.49-0.78)	<.01	0.71 (0.67-0.78)	<.01	0.51 (0.47-0.55)	<.001
Braden score	-	-	0.43 (0.27-0.58)	.053	0.45 (0.28-0.52)	.063	0.53 (0.34-0.58)	.074
Moisture subscore	-	-	0.38 (0.29-0.42)	.069	0.42 (0.27-0.48)	.053	0.39 (0.31-0.46)	.072
Sensitivity perception subscore	−1.78	0.239	0.34 (0.29-0.39)	<.001	0.31 (0.25-0.34)	<.001	0.33 (0.23-0.36)	<.001
Activity subscore	-	-	0.57 (0.39-0.65)	.073	0.54 (0.36-0.61)	.057	0.58 (0.51-0.62)	.064
Mobility subscore	−1.23	0.347	0.67 (0.69-0.77)	<.001	0.64 (0.61-0.68)	<.001	0.68 (0.65-0.71)	<.001
Nutrition subscore	-	-	0.58 (0.51-0.65)	.124	0.53 (0.47-0.58)	.102	0.61 (0.55-0.75)	.143
Friction subscore	-	-	0.84 (0.76-0.94)	.068	0.81 (0.75-0.85)	.048	0.86 (0.76-0.83)	.075
classification of the admission cause	-	-	0.95 (0.87-1.03)	.471	0.86 (0.84-0.95)	.036	1.21 (1.05-1.23)	.015
Surgical history	1.94	0.208	1.58 (1.17-1.76)	<.001	1.34 (1.30-1.53)	<.01	1.87 (1.79-1.94)	<.001

Abbreviations: BMI, body mass index; BUN, blood urea nitrogen; GCS, Glasgow Coma Scale; HCT, haematocrit; OR, odd ratio; PLT, platelet.

[a]The OR and *P* values represent the comparison between the total PI samples in ICU (N = 206) and the non-PI samples (N = 412).

[b]The OR and *P* values represent the comparison between the PI-developed samples within 24 hours in ICU (N = 53) and the non-PI samples (N = 412).

[c]The OR and *P* values represent the comparison between the PI-developed samples after 24 hours in ICU (N = 153) and the non-PI samples (N = 412).

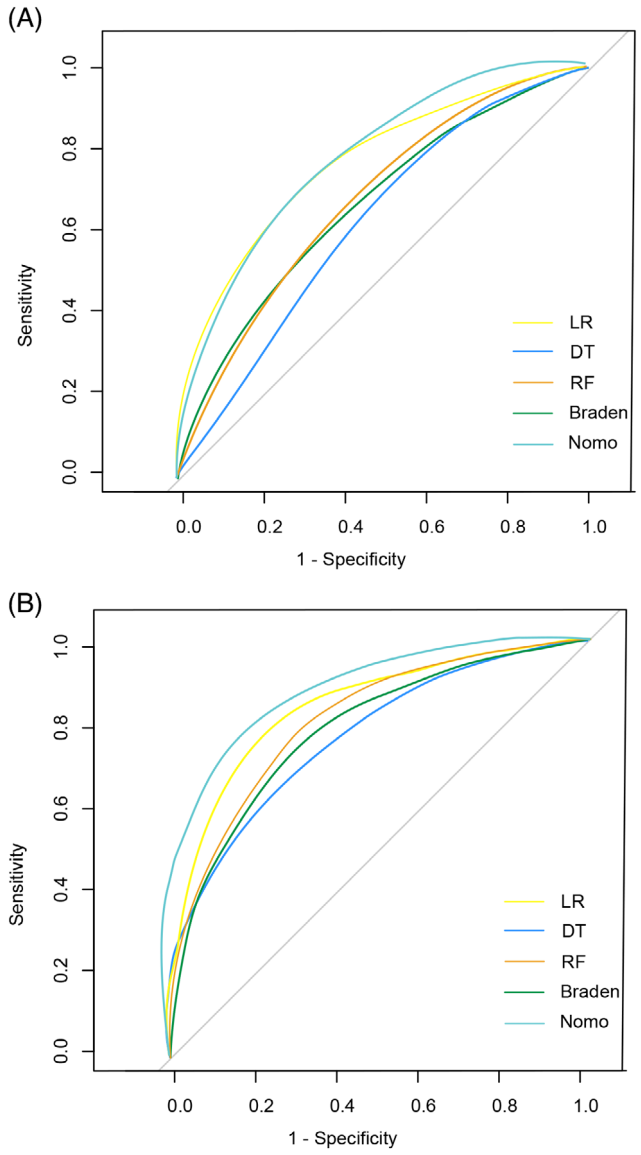

FIGURE 1 The comparison regarding area under receiver operating curves of the three machine learning-based clinical models, Braden scale and integrative nomogram in the training and test cohorts, respectively

3.4 | Nomogram construction and validation

According to the principle to achieve the highest sensitivity and a relative lower specificity above 0.8, a multivariate logistic regression model was picked out as the top-performing basic model with seven features shown in Table 5. We then incorporated the Braden score into the logistic regression model with seven significant parameters and constructed the integrative nomogram. The comparison of AUROC in Figure 1 revealed that there were obvious differences in performance between the nomogram and the basic machine learning models in both training and test cohort.

The prediction performance details on NRI and IDI were presented in Table 6. When compared with each other, a consistent trend was noticed that NRI or IDI favouring LR

TABLE 6 Prediciton value of the basic models and nomogram regarding net reclassification improvement (NRI) and integrated discrimination improvement (IDI)

		Variables	
Model	**Cohort**	**NRI**	**Absolute_IDI**
LR	Training	-	-
	Test	-	-
DT (vs LR)	Training	−0.17	−0.36
	Test	−0.12	−0.22
RF (vs LR)	Training	−0.34	−0.47
	Test	−0.28	−0.65
Nomogram (vs LR)	Training	0.43	0.51
	Test	0.37	0.39

Abbreviations: DT, decision tree; LR, logistic regression; RF, random forest.

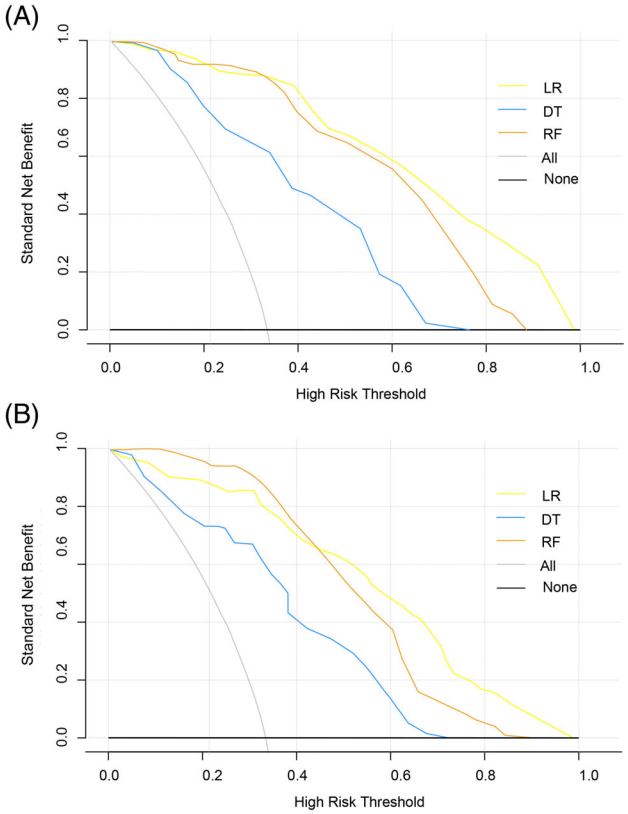

FIGURE 2 Decision curve analyses of the three machine learning-based clinical models, Braden scale and integrative nomogram in the training and test cohorts, respectively

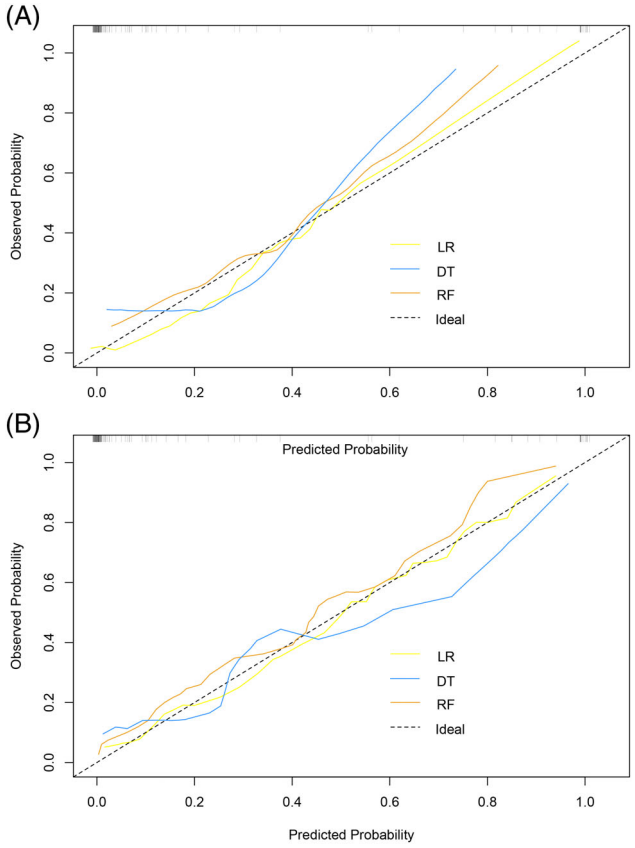

FIGURE 3 Calibration curves of the three machine learning-based clinical models, Braden scale and integrative nomogram in the training and test cohorts, respectively

model over the other basic models. Compared with three basic models (LR, DT and RF), the LR model had significant improvement in reclassification with NRI and IDI above 0. We performed a decision curve analysis (DCA) to assess the applicability of each model. The DCA exhibited the largest area under the decision curve, which meant preferable net benefits of the nomogram (Figure 2) in both training and test cohort than others. Calibration capacity shows the best concordances between the nomogram-based predictions and observed probabilities (Figure 3).

Overall, the nomogram integrating seven parameters identified in the machine learning and the classical Braden score was evidenced as the top-performing model for predicting the risk of PIs in ICU and was exhibited in Figure 4.

4 | DISCUSSIONS

To our best knowledge, this is the first study to exploit different machine learning algorithms to construct a seven-variable scoring nomogram on merit for assessment of the risk of PI in ICU. Braden scale was proposed in the nursing field to take preventive measures to arrest the PI development. However, extra focus on the Braden assessment item caused a nursing burden in the circumstances with tight medical resources.[24] Besides, the accessibility of the Braden scale index was not as convenient as the parameters shown in our clinical models, thus applicability and generalisation were not superior enough.[25]

Since the Braden scale was not perfect to use,[26] we performed the comparison between the Braden scale and the basic machine learning model and found that although discrimination capacity was comparable to some extent, there existed complementarity among the clinic- and bedside nursing-acquired parameters. Thus, a nomogram prognostic model comprising both clinical features, and the Braden scale was developed in this study.

In the preliminary work, once admitted into ICU, easy application of clinical parameters in LR model except for Braden score were proposed for risk stratification, followed by the Braden scoring specifically aiming at the high-risk subgroup to quantify the possibilities further. The combination of two assessment methods in chronological order displayed a dramatic increment of predictive power and resource allocation efficiency.

In the comprehensive nomogram, advanced age, bilirubin, peripheral vascular disease, diabetes mellitus, surgical history and GCS were added to the Braden scale, which was consistent with the previous analysis to a large extent.[27-30] A physiologic risk factor like bilirubin was presumed as a marker of hepatic and cystic abnormalities, accumulating inflammatory infiltration in the serum and in turn arousing inflammatory and oxidative responses.[22] The imbalance of immunity homeostasis dampened the output of the healing process and inclined toward a wounded outcome.

Patients with low GCS scores could exhibit a wide range of mobility-related disabilities. In accordance with peripheral vascular dysfunction, hemodynamic disorder owing to blood stasis decreased the perfusion, and devascularisation of surrounding tissues in high demand of blood supply led to skin damage. Previous surgical history was considered as a more influential risk factor, adding six folds chances at the incidence to patients experiencing surgery beyond those without surgery, shedding light on the physiological changes under the stress of surgery.

Given those PI events occurred in the general ward before ICU, their classification into non-PI developed in the ICU subgroup generated the false-negative rate. The second occurrence of PI was eclipsed by the average intensity of nursing prevention.

The subclassification partitioned by the period within or beyond 24 hours from ICU admission to PI attack

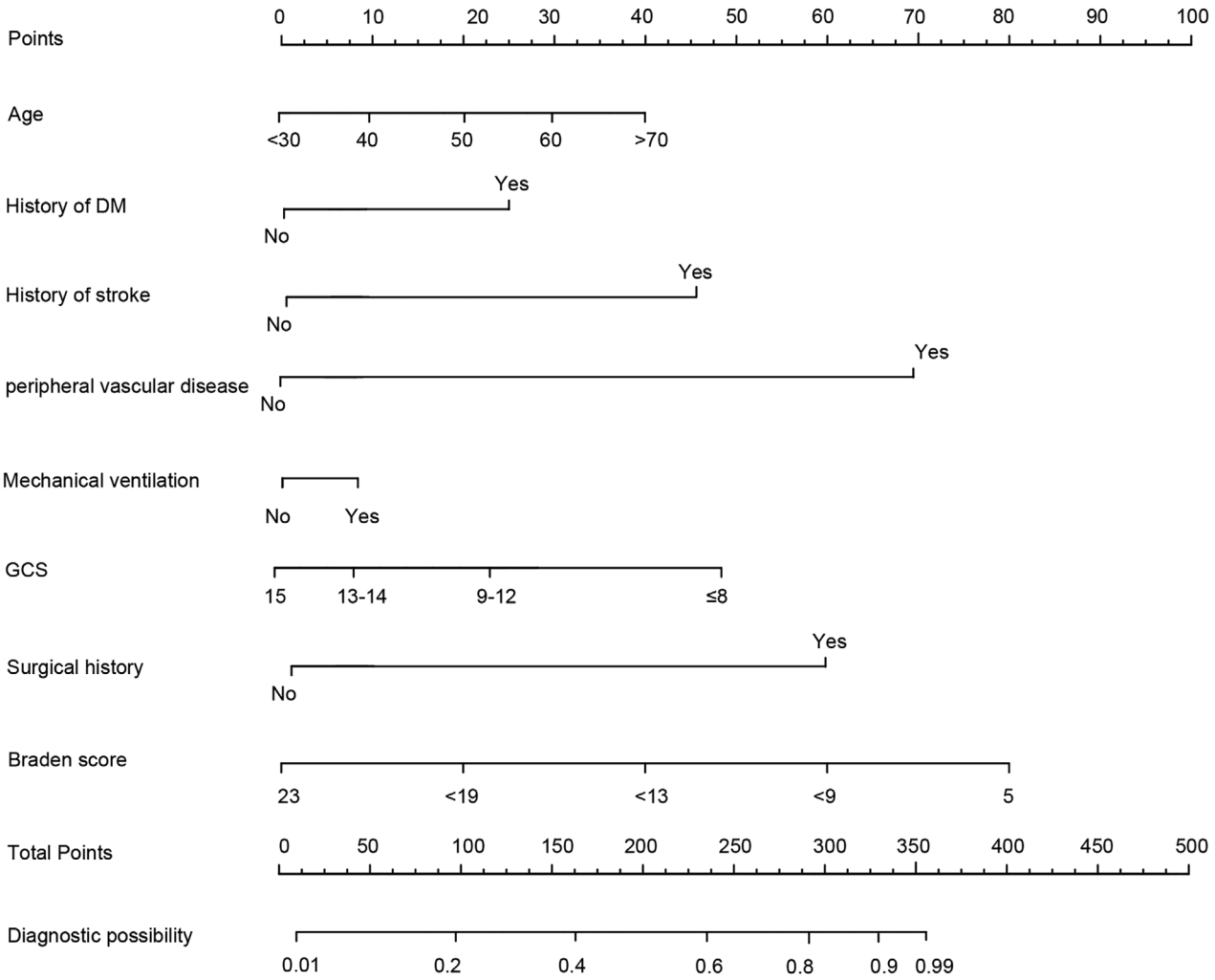

FIGURE 4 The nomogram for predicting the risk probability of PIs in ICU

reflected the disparity of the feature mechanism. The early-developed PIs always occurred in those who suffered a long hospital encounter with chronic consumptive diseases as evidenced by the OR of 0.86 favouring a chronic consumptive admission cause. The accumulation of malnutrition and the break of homeostasis burst out at the time point of analysis and provoked the PI outcome. On the other hand, acute critical diseases contributed to late-developed PIs. Injury markers less influential in the early-developed subgroup, such as blood urea nitrogen (BUN), played a majority role in the late-developed subgroup.

Given the previous incomplete understanding of causality in the development of PI, the novel addition of machine learning nourished the integrative competence of the basic clinical models.[31,32] Optimising the allocation of nursing resources liberated nursing strengths from the routine ineffective operation. In addition, the inclination to a lower threshold of specificity and a compensating high-sensitivity threshold signified the abandonment of centralised nursing resources. Equilibrium assignment of nursing preventions to every potential victim before PI developed, although seeming like a simplified and crude way, truly functioned in the refrainment of pressure injuries. Our clinical model orchestrated relatively complete clinical information in comparison with the previous models and was well equipped to identify more risk factors extensively connected with PI in ICU.

There still existed some limitations to declare. Albeit with the multicollinearity eliminated by the multivariate analysis, the existing interrelationship between clinical parameters obscured the causality. Another inescapable point lied in the hard extrapolation of our model to other medical institutions. The vigorous dependency of the feature extraction on the output port (including EHR and nursing labour) rendered this model by no means guaranteed to achieve a strong prognostic power. Furthermore, a sequential prospective cohort to test the model in the real world would be persuasive in favour of the predictive model.

5 | CONCLUSIONS

The novel nomogram was equipped with excellent performances regarding discrimination, calibration and clinical utility. Fusion of Braden scale and machine learning-based clinical model was validated to benefit the prevention of PI in ICU.

ACKNOWLEDGEMENTS

This study was supported by the Natural Science Foundation of Zhejiang Province (LQ17H260004); the Science and Technology Foundation of Wenzhou (Y20180038), Science and Technology Association of Wenzhou (rxk02).

CONFLICT OF INTEREST

The authors have no relevant affiliations or financial involvement with any organisation or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending or royalties.

AUTHOR CONTRIBUTIONS

Jie Xu and Danxiang Chen designed this study, collected and analysed data and drafted the manuscript; Caixia Sun designed this study, revised the manuscript and supervised this study; Xiaoyun Pan and Yu Chen collected and analysed the data; Xiaoming Zhuang analysed the data, interpreted the results and reviewed the manuscript. All authors read and approved the final manuscript.

DATA AVAILABILITY STATEMENT

The clinical data supporting the conclusions of this study were available on the electronic health records of the First Affiliated Hospital of Wenzhou Medical University, and access can be provided upon request to the authors.

REFERENCES

1. Mervis JS, Phillips TJ. Pressure ulcers: pathophysiology, epidemiology, risk factors, and presentation. *J Am Acad Dermatol*. 2019;81(4):881-890.
2. VanGilder C, Amlung S, Harrison P, Meyer S. Results of the 2008-2009 international pressure ulcer prevalence survey and a 3-year, acute care, unit-specific analysis. *Ostomy Wound Manage*. 2009;55(11):39-45.
3. Labeau SO, Afonso EP, Llaurado-Serra M. Pressure injury research in the ICU: getting rid of a black spot. *Crit Care Med*. 2019;47(2):e161–e162.
4. Cox J, Schallom M, Jung C. Identifying risk factors for pressure injury in adult critical care patients. *Am J Crit Care*. 2020;29(3):204-213.
5. Alderden J, Drake KP, Wilson A, Dimas J, Cummins MR, Yap TL. Hospital acquired pressure injury prediction in surgical critical care patients. *BMC Med Inform Decis Mak*. 2021;21(1):12.
6. Tervo-Heikkinen TA, Heikkila A, Koivunen M, et al. Pressure injury prevalence and incidence in acute inpatient care and related risk factors: a cross-sectional national study. *Int Wound J*. 2021. http://dx.doi.org/10.1111/iwj.13692
7. Alshahrani B, Sim J, Middleton R. Nursing interventions for pressure injury prevention among critically ill patients: a systematic review. *J Clin Nurs*. 2021;30(15–16):2151-2168.
8. Razmus IS, Keep SM. Neonatal intensive care nursing pressure injury prevention practices: a descriptive survey. *J Wound Ostomy Continence Nurs*. 2021;48(5):394-402.
9. Huang C, Ma Y, Wang C, et al. Predictive validity of the Braden scale for pressure injury risk assessment in adults: a systematic review and meta-analysis. *Nurs Open*. 2021;8(5):2194-2207.
10. Curley MAQ, Hasbani NR, Quigley SM, et al. Predicting pressure injury risk in pediatric patients: the Braden QD scale. *J Pediatr*. 2018;192(189–195):e182.
11. Higgins J, Casey S, Taylor E, Wilson R, Halcomb P. Comparing the Braden and Jackson/Cubbin pressure injury risk scales in trauma-surgery ICU patients. *Crit Care Nurse*. 2020;40(6):52-61.
12. Jin Y, Jin T, Lee SM. Automated pressure injury risk assessment system incorporated into an electronic health record system. *Nurs Res*. 2017;66(6):462-472.
13. Xu W, Yu X, Jiang H, Wang Y, Ye Y. A nomogram prediction of pressure injury in critical ill patients: a retrospective cohort study. *Int Wound J*. 2021. http://dx.doi.org/10.1111/iwj.13680
14. Cheng H, Sun X, Ji X, et al. Risk factors and the potential of nomogram for predicting hospital-acquired pressure injuries. *Int Wound J*. 2020;17(4):974-986.
15. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-349.
16. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet*. 2011;43(1):27-33.
17. Hu YH, Lee YL, Kang MF, Lee PJ. Constructing inpatient pressure injury prediction models using machine learning techniques. *Comput Inform Nurs*. 2020;38(8):415-423.
18. Song W, Kang MJ, Zhang L, et al. Predicting pressure injury using nursing assessment phenotypes and machine learning methods. *J Am Med Inform Assoc*. 2021;28(4):759-765.
19. Zhao RF, Zhang WY, Zhou L, Chen Y. Building a predictive model for successful vaginal delivery in nulliparas with term cephalic singleton pregnancies using decision tree analysis. *J Obstet Gynaecol Res*. 2019;45(8):1536-1544.
20. Zhang ZQ, He G, Luo ZW, et al. Individual mortality risk predictive system of patients with acute-on-chronic liver failure based on a random survival forest model. *Chin Med J*. 2021;134(14):1701-1708.
21. Huang JC, Tsai YC, Wu PY, et al. Predictive modeling of blood pressure during hemodialysis: a comparison of linear model, random forest, support vector regression, XGBoost, LASSO regression and ensemble method. *Comput Methods Prog Biomed*. 2020;195:105536.
22. Khlifi L, Graiet H, Sahli S, et al. Evidence of metabolic imbalance and oxidative stress among patients suffering from pressure ulcers. *J Dermatolog Treat*. 2019;30(4):414-421.
23. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak*. 2006;26(6):565-574.
24. Cremasco MF, Wenzel F, Zanei SS, Whitaker IY. Pressure ulcers in the intensive care unit: the relationship between

nursing workload, illness severity and pressure ulcer risk. *J Clin Nurs*. 2013;22(15–16):2183-2191.

25. Garcia-Fernandez FP, Pancorbo-Hidalgo PL, Agreda JJ. Predictive capacity of risk assessment scales and clinical judgment for pressure ulcers: a meta-analysis. *J Wound Ostomy Continence Nurs*. 2014;41(1):24-34.
26. Wei M, Wu L, Chen Y, Fu Q, Chen W, Yang D. Predictive validity of the Braden scale for pressure ulcer risk in critical care: a meta-analysis. *Nurs Crit Care*. 2020;25(3):165-170.
27. Lyder CH, Wang Y, Metersky M, et al. Hospital-acquired pressure ulcers: results from the national Medicare patient safety monitoring system study. *J Am Geriatr Soc*. 2012;60(9):1603-1608.
28. Sprigle S, McNair D, Sonenblum S. Pressure ulcer risk factors in persons with mobility-related disabilities. *Adv Skin Wound Care*. 2020;33(3):146-154.
29. Coyer F, Tayyib N. Risk factors for pressure injury development in critically ill patients in the intensive care unit: a systematic review protocol. *Syst Rev*. 2017;6(1):58.
30. Tayyib N, Coyer F, Lewis P. Saudi Arabian adult intensive care unit pressure ulcer incidence and risk factors: a prospective cohort study. *Int Wound J*. 2016;13(5):912-919.
31. Jablonski AM, DuPen AR, Ersek M. The use of algorithms in assessing and managing persistent pain in older adults. *Am J Nurs*. 2011;111(3):34-43; quiz 44–35.
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*. 2019;7(1):49.

How to cite this article: Xu J, Chen D, Deng X, et al. Development and validation of a machine learning algorithm–based risk prediction model of pressure injury in the intensive care unit. *Int Wound J*. 2022;19(7):1637-1649. doi:10.1111/iwj.13764