Original Research Article

DIGITAL HEALTH
Volume 10: 1–10
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241277030
journals.sagepub.com/home/dhj

Machine learning-based prediction of coronary care unit readmission: A multihospital validation study

Fei-Fei Flora Yau[1], I-Min Chiu[1,2], Kuan-Han Wu[1], Chi-Yung Cheng[1], Wei-Chieh Lee[3], Huang-Chung Chen[4], Cheng-I Cheng[4] and Tien-Yu Chen[4]

Abstract

Objective: Readmission to the coronary care unit (CCU) has significant implications for patient outcomes and healthcare expenditure, emphasizing the urgency to accurately identify patients at high readmission risk. This study aims to construct and externally validate a predictive model for CCU readmission using machine learning (ML) algorithms across multiple hospitals.

Methods: Patient information, including demographics, medical history, and laboratory test results were collected from electronic health record system and contributed to a total of 40 features. Five ML models: logistic regression, random forest, support vector machine, gradient boosting, and multilayer perceptron were employed to estimate the readmission risk.

Results: The gradient boosting model was selected demonstrated superior performance with an area under the receiver operating characteristic curve (AUC) of 0.887 in the internal validation set. Further external validation in hold-out test set and three other medical centers upheld the model's robustness with consistent high AUCs, ranging from 0.852 to 0.879.

Conclusion: The results endorse the integration of ML algorithms in healthcare to enhance patient risk stratification, potentially optimizing clinical interventions, and diminishing the burden of CCU readmissions.

Keywords

Multihospital validation, machine learning, coronary care unit, readmission, gradient boosting

Submission date: 13 April 2024; Acceptance date: 6 August 2024

Introduction

Readmission to coronary care unit (CCU) after initial hospitalization is a critical problem that has significant implications for patient outcomes and healthcare costs. Patients who are readmitted to intensive care unit (ICU) have worse outcomes, including higher mortality rates, longer hospital stays, and increased healthcare costs.[1,2] According to previous studies, the readmission rate to CCU ranges from 2.5% to 8% before discharge of initial hospitalization.[3–5]

Identifying patients at high risk of readmission is essential for improving patient outcomes and reducing healthcare costs. By identifying high-risk patients, healthcare

[1]Department of Emergency Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan
[2]Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[3]Division of Cardiology, Department of Internal Medicine, Chi Mei Medical Center, Tainan, Taiwan
[4]Division of Cardiology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan

Corresponding author:
Tien-Yu Chen, Division of Cardiology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, No. 123, Dapi Rd., Niaosong Dist., Kaohsiung City, 83301, Taiwan.
Email: a_den0904@hotmail.com

providers can implement targeted interventions, such as medication adjustments, lifestyle modifications, and care coordination, to prevent readmission and improve outcomes. Therefore, developing accurate predictive models for readmission risk is critical for improving patient care and reducing healthcare costs.

Machine learning (ML) has emerged as a powerful tool in healthcare for predicting various outcomes and improving patient care. ML algorithms can detect complex patterns and relationships in data that may not be apparent to humans, making it a valuable tool for predicting readmission after discharge from CCU. Previous studies have used ML to predict readmission risk in various patient populations, including those with heart failure, chronic obstructive pulmonary disease, and diabetes.[6–8] Huberts et al.[9] demonstrated the effectiveness of ML algorithms in predicting unplanned readmission and mortality in cardiovascular patients, highlighting key risk factors and achieving high predictive performance in their study. Similarly, Rojas et al.[10] developed an ML model to predict ICU readmission, showing significantly better performance than previously published algorithms, underscoring the potential of ML in improving patient outcomes in critical care settings.

The objective of this study is to develop a predictive model for readmission to CCU after initial hospitalization using ML algorithms (Graphic Abstract). Specifically, we aim to identify patient characteristics and clinical factors that are associated with readmission and develop a model that can accurately predict readmission risk.

Method

Study setting

In this prognostic study, we leveraged electronic health records from Lin Kou Chang Gung Memorial Hospital (derivation cohort) collected from 1 October 2000 to 30 June 2019, to establish the development cohort. For external validation, we included records from three major medical centers located diversely nationwide. These centers are Kaohsiung Chang Gung Hospital (validation cohort A), Chiayi Chang Gung Hospital (validation cohort B), and Keelung Chang Gung Hospital (validation cohort C). The data spanned from 1 January 2001 to 30 June 2019. All four hospitals are recognized as the largest healthcare facilities in their respective local areas, ensuring a diverse and comprehensive dataset for this study. Patients' informed consent was waived due to the retrospective nature of the study using de-identified electronic health record data. Ethical approval was obtained from the Institutional Review Board of Chang Gung Medical Foundation. The approval reference number is 202201541B0.

Data collection

In the development cohort, we included all patients admitted to the CCU during the study period. We excluded patients who died or were transferred to other hospitals during their CCU admission. The complete dataset was randomly partitioned into three subsets at an $8:1:1$ ratio, designated as the training, internal validation, and hold-out test sets, respectively. To ensure the integrity of the data, we used patient-level splitting; this approach guaranteed that different admissions of the same patient were not included in multiple subsets (Figure 1).

For this study, we collected a total of 40 features from the Electronic Health Records (EHR), encompassing patient demographic characteristics, laboratory test results, clinical diagnosis, medical history, procedure and clinical event occurred in CCU admission, all of which are relevant to the prognosis during the CCU admission. Diagnostic information upon and during CCU admission, as well as cardiovascular treatments, were also gathered as predictive features (see Table 1). The primary outcome was CCU readmission, defined as a return to the CCU or death occurring postdischarge from the CCU within the same hospital stay.

For the laboratory tests, we gathered the standard tests, such as complete blood count, renal function, and electrolyte conducted during the same admission and selected the test closest to the CCU discharge for analysis. For the left ventricular ejection fraction, obtained from echocardiography reports within 1 year, we used the measurement nearest to the CCU discharge. Information on procedures and events was collected and identified in relation to the same episode of CCU admission. The statistical comparison of collected features in readmission and control group were demonstrated in Supplemental Table 2.

Missing values were observed in the numerical features, with missing rates ranging between 1.2% and 4.5%. These missing values were imputed using the median value from the remaining data. Before model training, all feature values were normalized to a scale of [0,1]. Numerical features were scaled using min–max normalization, which adjusts each feature value by subtracting the minimum value of the feature and dividing by the range of the feature values. Categorical features were encoded using one-hot encoding to transform them into a format suitable for ML algorithms.

Development of machine learning models

In this study, we employed three ML models, namely logistic regression (LR), random forest (RF), support vector machine (SVM), gradient boosting (GB), and multilayer perceptron (MLP) to predict the risk of patients being readmitted to the CCU. LR is a fundamental statistical model widely employed in ML for binary classification tasks.[11–13] By examining a cohort's clinical predictors, the LR model estimates the probability of readmission by assigning weights to input features. The model's output is a value between 0 and 1, interpreted as the likelihood of

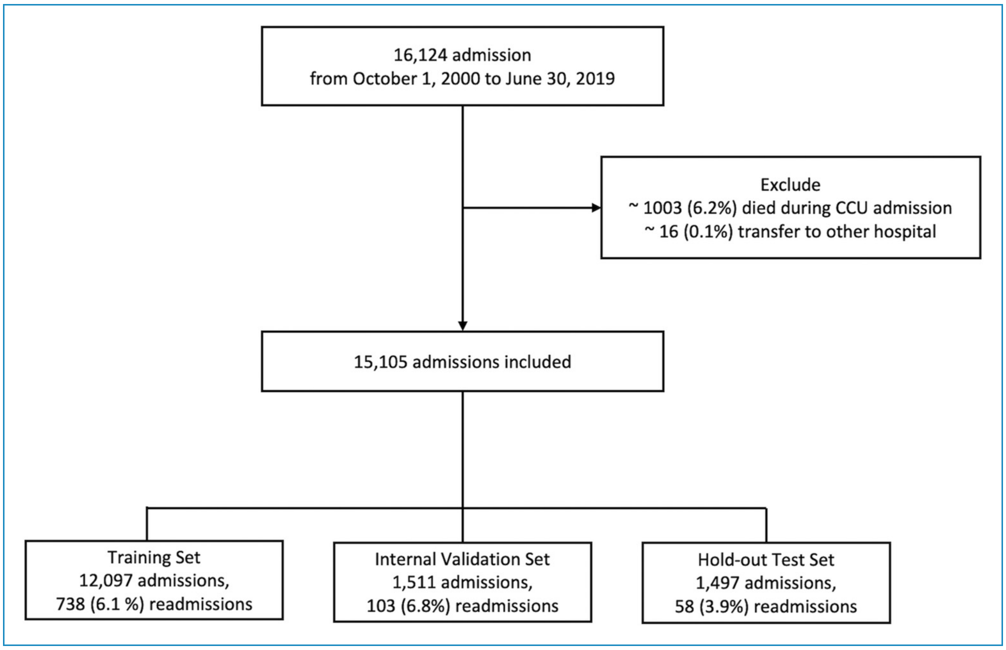

Figure 1. Patient inclusion flowchart of this study.

Table 1. Patient's demographics of included hospitals.

Variable	Derivation cohort n = 15,105 N(%)/mean(SD)	Validation cohort A n = 9651 N(%)/mean(SD)	Validation cohort B n = 8821 N(%)/mean(SD)	Validation cohort Cn = 6608 N(%)/mean(SD)
Sex				
Male	9796(64.9%)	6715(69.6%)	5378(61%)	4024(60.9%)
Female	5309(35.1%)	2936(30.4%)	3443(39%)	2584(39.1%)
Age	66.2(14.6)	65.9(13.5)	70.7(14)	69.4(14.6)
Main diagnosis				
ACS	9158(60.6%)	6640(68.8%)	4466(50.6%)	3385(51.2%)
OHCA	77(0.5%)	50(0.5%)	156(1.7%)	23(0.3%)
IHCA	643(4.2%)	567(5.9%)	465(5.3%)	381(5.8%)
Readmission	899(6%)	520(5.4%)	569(6.5%)	409(6.2%)

ACS: acute coronary syndrome; IHCA: in-hospital cardiac arrest; OHCA: out-of-hospital cardiac arrest.

an event occurring, in this study, the return of a patient to CCU. This method is particularly valued for its simplicity, making it a baseline standard tool for predictive analytics. RF is a type of ensemble learning algorithm that is based on the decision tree method. It employs bagging, a sampling technique that randomly selects subsets of the training data and features to create multiple classification and regression trees. The results of the individual trees are then aggregated through a voting process to arrive at the final prediction. The idea behind this technique is that each tree is slightly different due to the random selection of data and features, which reduces the risk of overfitting to the training data.

In various medical studies, random forest has shown higher predictive power that has been widely adopted for predicting risk outcomes and predictive factors.[14–16] SVM works by finding the optimal hyperplane that best separates the data into different classes. The hyperplane is chosen to maximize the margin between the classes, ensuring the most significant possible distance between the closest points of the classes (support vectors).[17] This technique is beneficial for high-dimensional spaces and cases where the number of dimensions exceeds the number of samples. SVMs are known for their robustness and effectiveness in scenarios where a clear margin of separation exists, and they have been successfully applied in various medical predictive analytics for their precision and generalization capabilities.[18] GB operates by sequentially building an ensemble of weak learners, each one focusing on correcting the errors of its predecessor. Each weak learner, typically a decision tree, is added to the ensemble with the goal of reducing the loss function, a measure of how well the model fits the entire dataset.[19] The performance of each tree in the sequence is taken into account by optimizing an objective function that combines the predictions with the actual values. This technique often results in a powerful predictive model that can handle various types of data and different distributions. GB has been widely recognized for its effectiveness in classification and regression tasks, often achieving high accuracy and robustness in predictive analytics challenges. MLP is a class of feedforward artificial neural network that consists of multiple layers of neurons. Each neuron in the network represents a computational unit that processes input data through a nonlinear activation function. The MLP learns to map inputs to outputs by adjusting the weights through a process called backpropagation, which minimizes the error between predicted and actual outcomes. MLP excels at capturing complex patterns in data and has been widely used in predictive analytics for its ability to model nonlinear relationships and interactions among features, leading to highly accurate predictions in clinical outcome studies.[20,21]

In the development of our ML models, we employed a grid search approach for hyperparameter tuning, using values that are commonly applied in practice. The goal of the grid search was to achieve the best performance in terms of the area under the receiver operating characteristic curve (AUC) on the internal validation set, which was also used to compare model performance. For LR, the best hyperparameter identified was a regularization parameter (C) set to 1.0. For RF, the optimal configuration included 100 trees (n_estimators), with the minimum number of samples required to split a node (min_samples_split) set to 2, and the minimum number of samples required at each leaf node (min_samples_leaf) set to 1. For the SVM, the grid search determined that the regularization parameter (C) should be 1.0, and the kernel type should be a radial basis function. For GB, the model was best configured with 100 boosting stages (n_estimators), a learning rate of 0.1 to control the contribution of each tree, and the maximum depth of individual regression estimators set to 3. Additionally, the minimum number of samples required to split a node was set to 2. The MLP, the optimal setup included one hidden layer of 100 neurons, using the ReLU activation function for nonlinearity between layers. The learning rate was set to "adaptive," allowing it to adjust as training progressed to minimize loss. The regularization parameter (alpha) was set to 0.0001 to help prevent overfitting by penalizing large weights.

We incorporate feature selection process during training ML model in this study. We adopted a forward stepwise feature selection algorithm that helps to select the optimal subset of features for each ML models. Forward stepwise feature selection is a process used in model building that begins with an empty model and adds in features one by one. In each step, the feature that provides the most significant improvement to the model fit is included, until the addition of new feature no longer significantly improves model performance.[22] The rationale behind choosing the forward stepwise feature selection method is multifactorial. First, it allows for a deliberate escalation in model complexity and ensures that each feature incorporated plays a pivotal role in the predictive accuracy of the model. Second, it prevents overfitting thereby bolstering the model's generalizability, and it augments interpretability by focusing only on the most important features. This feature selection method is particularly useful when dealing with datasets with a large number of variables.

In this study, we utilized an 8 : 1 : 1 dataset split for the training, internal validation, and hold-out test sets, respectively. The forward stepwise feature selection process involved the following steps: starting with an empty model, evaluating all candidate features not yet included in the model at each iteration, and adding the feature that most improves model performance as measured by the AUC in the internal validation set. This process continued until no additional feature significantly improved the model performance. The whole feature selection process was validated on a separate validation set, ensuring that the selected features contributed to the model's performance without overfitting.

Statistical analysis

Continuous variables are expressed as mean ± standard deviation, while nominal variables are presented as proportions. Ordinal variables are depicted as median along with their interquartile range or mean (standard deviation) if they were normally distributed. The comparative performance of five ML algorithms, LR, RF, SVM, GB and MLP was assessed with the AUC and F-1 score serving as the primary evaluation metric in the interval validation set. Upon selecting the optimal model, its robustness was further evaluated in the hold-out test set and also three different medical centers as external validation, with performance metrics including sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). All analyses were conducted using Python 3.8, utilizing the Scikit-Learn package.

Results

The demographic data for all included hospitals are presented in Table 1. Our model was trained using data from derivation cohort, which had the highest patient count of 15,105, followed by validation cohort A with 9651 patients, validation cohort B with 8,821, and validation cohort C with 6608. In derivation cohort, there was a male predominance of 64.9% (9796 patients), and the mean age was 66.2 ± 14.6 years. The primary diagnosis was acute coronary syndrome (ACS), accounting for 60.6% of cases, and the readmission rate was 6%. For the three medical centers used as external validation, the proportion of male patients ranged from 60.9% to 69.6%, the mean patient age varied from 65.9 to 70.7 years. The prevalence of ACS was higher in validation cohort A (68.8%) compared to validation cohorts B (50.6%) and C (51.2%). Other diagnoses, such as sepsis, gastrointestinal bleeding, and chronic obstructive pulmonary disease (COPD), also showed variability. For instance, the incidence of sepsis was higher in validation cohorts B (35.2%) and C (31.6%) compared to the development cohort (19.9%). Procedural interventions during admission indicated variations as well. The use of intra-aortic balloon pump (IABP) was 5.7% in the development cohort, while it was higher at 14.7% in validation cohort A and lower at 4.7% in cohort B and 2.9% in cohort C. Extracorporeal membrane oxygenation (ECMO), a critical intervention for patients with severe cardiac and respiratory failure, was utilized in 4.1% of the development cohort. The usage varied significantly across the validation cohorts: 4.2% in validation cohort A, 1% in validation cohort B, and 1.7% in validation cohort C. Mechanical ventilation was used at similar rates: 11% in the development cohort, 11.9% in validation cohort A, 11.5% in validation cohort B, and 13.8% in validation cohort C. These variations in procedural interventions reflect differences in clinical practices and patient severity across the hospitals. This diversity in the datasets is beneficial for testing the robustness of the model, ensuring that it can generalize well across different clinical settings. Finally, the readmission rates across these hospitals varied from 5.4% to 6.5%. The complete list of variables collected is shown in Supplemental Table 1.

The outcome of the forward stepwise feature selection algorithm for determining the optimal subset of predictive features for CCU readmission is delineated in Table 2. Across the three ML models analyzed, the use of a ventilator and the length of CCU stay were consistently selected as the first and second most significant features, respectively. For the LR model, platelet count, COPD, sodium levels, IABP use, in-hospital cardiac arrest, and creatinine levels were identified as the subsequent predictors in order of importance. The GB model included additional predictors such as hypertension, peripheral artery occlusive disease, myocardial infarction (MI), ECMO usage, coronary artery disease, diabetes mellitus (DM), and atrial fibrillation. The RF model identified COPD, ECMO, and out-of-hospital cardiac arrest as other important features. Notably, certain predictors like IABP and COPD were selected by multiple models across different modeling approaches.

The performance of the five models, as per their respective AUCs, was commendable with scores of 0.858, 0.875, 0.887, 0.877, and 0.862 in the internal validation set, for LR, RF, GB, MLP, and SVM, respectively (Figure 2a). Correspondingly, their F-1 scores were 0.253 for LR, 0.271 for RF, 0.284 for GB, 0.271 for MLP, and 0.255 for SVM. Accounting for performance, the GB model was chosen to predict the risk of CCU readmission in hold-out test set and external validation. In this study, external validations were conducted on patient data procured from diverse geographical regions. The model exhibited significant discriminatory ability in predicting CCU readmissions, supported by consistent high AUCs in test set of 0.879, and from validation cohort A, B, and C, with values of 0.863, 0.848, and 0.852, respectively, as illustrated in Figure 2b.

For binary classification, the default decision threshold is set to 0.5. However, this could be adjusted to substantially impact the model's performance owing to the trade-off between false positives and false negatives. Table 3 demonstrates the model's performance in predicting CCU readmission at different cutoff values within the hold-out test set and external validation sets from validation cohort A, B, and C. The threshold of 0.025, identified by the Youden Index as optimal, yields a balance with a sensitivity of 0.915 and specificity of 0.724 in the hold-out test set, accompanied by a PPV of 0.154 and NPV of 0.995. The selection of cutoff values at 0.125 is strategic, aiming to attain a specificity of 0.9 in the hold-out test set. At a cutoff value of 0.125, specificity is prioritized, while sensitivity decreases to 0.534, and PPV rises to 0.233, and NPV of 0.969. The respective sensitivity, specificity, PPV, and NPV to three external validation hospitals were demonstrated in Table 3.

In this study, we adopted Shapley Additive Explanation (SHAP) algorithm for model interpretation, which provides a unified measure of feature importance and allows for the evaluation of each feature's contribution to the model's prediction. The SHAP value analysis for the GB model predicting CCU readmission, was illustrated in Figure 3. The bar plot (Figure 3a), shows that the length of CCU stay is the most critical feature, contributing the most substantial

Table 2. Result variables of stepwise features selection.

	Logistic regression	Gradient boosting	Random forest	Multilayer perceptron	Support vector machine
1st feature	Ventilator	Ventilator	Ventilator	Length of CCU stay	Length of CCU stay
2nd feature	Length of CCU stay	Length of CCU stay	Length of CCU stay	Ventilator	Ventilator
3rd feature	Platelet	Hypertension	COPD	IABP	IABP
4th feature	COPD	PAOD	ECMO	eGFR	Creatinine
5th feature	Sodium	MI	OHCA	CAD	IHCA
6th feature	IABP	IABP		Male	Age
7th feature	IHCA	COPD		Platelet	LVEF
8th feature	Creatinine	ECMO		WBC	Na
9th feature		CAD		Age	CLD
10th feature		DM		Na	Dialysis
11th feature		Atrial fibrillation			
12th feature		OHCA			

CAD: coronary artery disease; CCU: coronary care unit; CLD: chronic lung disease; COPD: chronic obstructive pulmonary disease; DM: diabetes mellitus; ECMO: extracorporeal membrane oxygenation; eGFR: estimated glomerular filtration rate; IABP: intra-aortic balloon pump; IHCA: in-hospital cardiac arrest; LVEF: left ventricular ejection fraction; MI: myocardial infarction; OHCA: out-of-hospital cardiac arrest; PAOD: peripheral arterial occlusive disease; WBC: white blood count.

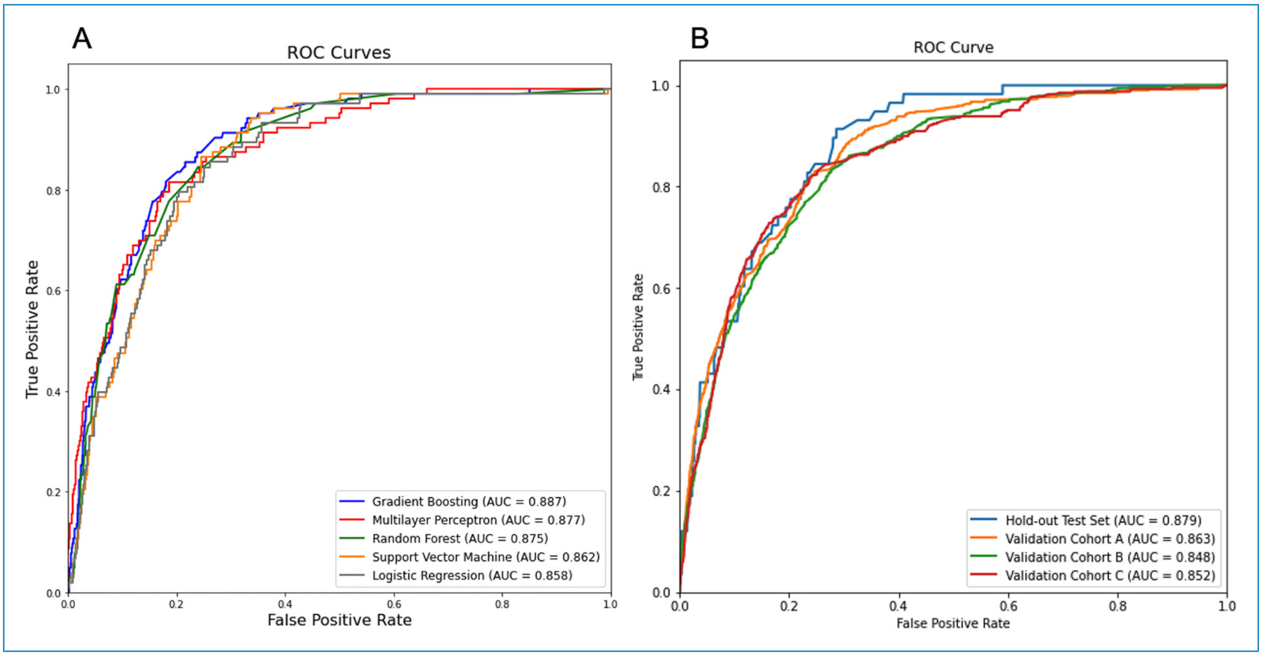

Figure 2. ROC curve and AUC value of (a) different machine learning models in internal validation and (b) gradient boosting in hold-out test set and external validation (b). AUC: area under the receiver operating characteristic curve; ROC: receiver-operating characteristic.

impact on the model's output. Other important features include the use of a ventilator, presence of MI, DM, hypertension, and atrial fibrillation. The dot plot (Figure 3b), elucidates the distribution and impact of SHAP values for each feature, illustrating the direction of influence. For instance, a longer CCU stay, ventilator use, presence of MI and atrial fibrillation are associated with higher readmission risk, as indicated by positive SHAP values. Conversely, features like DM and hypertension show mixed influences, with some instances contributing positively and others negatively to the readmission prediction.

Discussion

The development of a ML model for predicting the risk of readmission to the CCU before discharge from the initial hospitalization represents a significant advancement in healthcare practice. The study's findings suggest that when ML algorithms are properly trained and validated, they can provide an accurate and reliable method for identifying patients at risk of readmission (Graphic Abstract). ML is particularly adept at analyzing complex patterns and relationships in data, which makes it an ideal tool for predicting the risk of readmission. In this study, five ML models—LR, RF, SVM, GB, and MLP—performed exceptionally well, with AUCs ranging from 0.858 to 0.887. These results underscore the potential of ML in healthcare, especially in predicting readmission risk and identifying high-risk patients. Furthermore, we established two different thresholds in this study, which prioritize either high sensitivity or high specificity. By utilizing these thresholds, healthcare providers can implement targeted interventions, such as continuous monitoring in a general ward or intensive medical treatment, to mitigate the risk of CCU readmission, thereby improving patient outcomes and reducing healthcare costs.

ML algorithms can use a variety of data sources to predict readmission risk, including electronic health records, clinical notes, and demographic data. ML algorithms can also incorporate nontraditional data sources, such as social determinants of health, to improve the accuracy of predictions.[23] By incorporating multiple data sources, ML algorithms can develop more accurate predictive models and identify high-risk patients who may benefit from targeted interventions to prevent readmission. However, there are also challenges associated with using ML in healthcare. One major challenge is the need for high-quality data. ML algorithms require large amounts of data to develop accurate predictive models, and the quality of the data is critical for ensuring that the models are reliable and valid. In addition, there are concerns about bias in ML algorithms, as they may perpetuate existing disparities in healthcare if they are trained on biased data.[24]

Predicting the trajectory of patients recovering after leaving CCU is challenging. A 2-year study found that a

Table 3. Model performance on predicting CCU readmission in different cut-off value.

Cut off value	Hold-out test set				Validation cohort A				Validation cohort B				Validation cohort C			
	Sensitivity	*Specificity*	*PPV*	*NPV*	*Sensitivity*	*Specificity*	*PPV*	*NPV*	*Sensitivity*	*Specificity*	*PPV*	*NPV*	*Sensitivity*	*Specificity*	*PPV*	*NPV*
0.025*	0.915	0.724	0.154	0.995	0.819	0.759	0.162	0.987	0.926	0.555	0.126	0.991	0.936	0.495	0.110	0.992
0.125	0.534	0.900	0.233	0.969	0.514	0.923	0.277	0.971	0.645	0.852	0.231	0.972	0.741	0.823	0.216	0.980

*The optimal threshold was determined using the Youden Index.

CCU: coronary care unit; NPV: negative predictive value; PPV: positive predictive value.

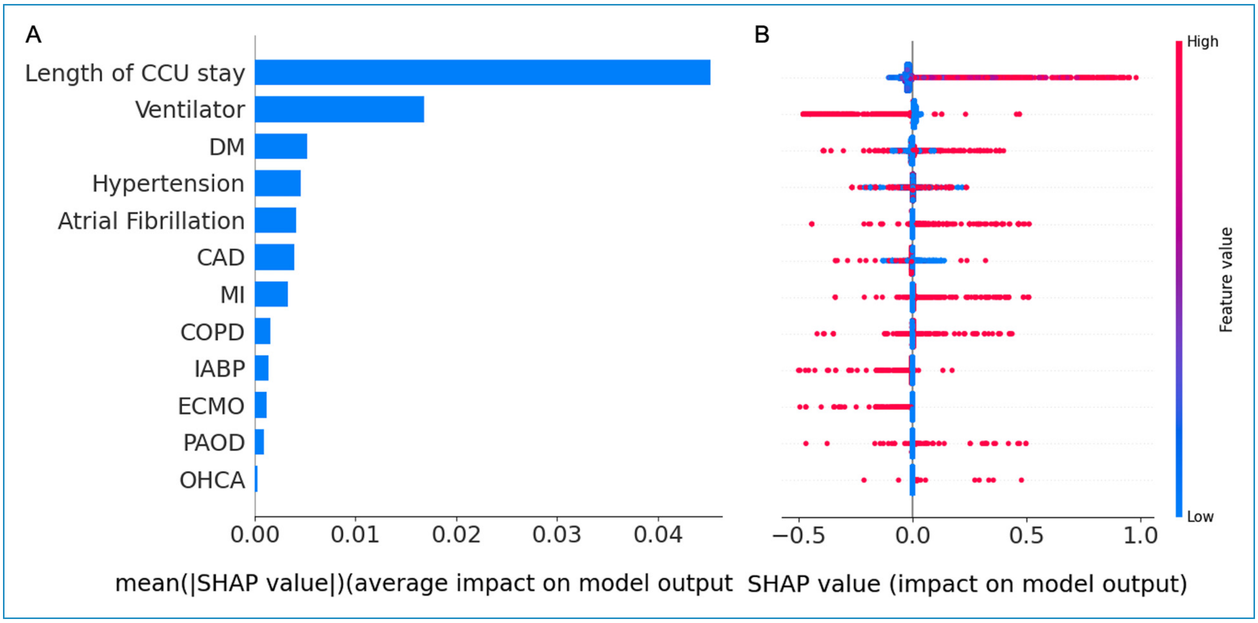

Figure 3. Interpretation of feature importance and contribution by SHAP value. SHAP: Shapley Additive Explanation.

quarter of patients experienced mortality after readmission to the CCU,[25] highlighting the importance of early post-CCU monitoring. However, monitoring every patient discharged from the CCU in the general ward can lead to excessive medical costs and increased burdens on healthcare workers. Thus, identifying patients at high risk of readmission to CCU during their hospital stay is critical. Huberts et al.[9] demonstrated the effectiveness of ML algorithms in predicting unplanned readmission and mortality in cardiovascular patients, highlighting key risk factors and achieving high predictive performance in their study. Similarly, Rojas et al.[10] developed an ML model to predict ICU readmission, showing significantly better performance than previously published algorithms, which underscores the potential of ML in improving patient outcomes in critical care settings. In our previous work, we introduced a risk stratify model to predict readmission to CCU with AUC range from 0.70 to 0.72.[3] Recently, one study using 14 features to develop a recurrent neural network to predict CCU readmission and found the balanced accuracy range from 0.619 to 0.727.[26] Another study utilizes 41 variables to develop ML model and had performance range from 0.744 to 0.986.[27] It appears that by having more associated features, we are able to build better model on it. Also, in this study we included all common risk factors, or features, found to be associated with CCU readmission across previous studies, including age, admission diagnosis, cardiac and noncardiac comorbidities, and laboratory test.[28–30] Through thorough data curation, we were able to include the interventions during admission to create a more comprehensive feature set.

In this study, we developed an ML model from an initial set of 40 variables and refined this list to select the best subset of features through stepwise feature selection while training the ML models. The stepwise iteration process was conducted independently for each model, resulting in a tailored set of predictors for each. For instance, while ventilator use and length of CCU stay were consistently identified as top features across all models, other variables like platelet count and hypertension were selectively included based on the specific model's algorithmic requirements and performance metrics. Through this approach, we were able to enhance the performance of each model by concentrating on the most informative features, thereby increasing the models' predictive accuracy and robustness. The resulting subsets of features are expected to provide the basis for more streamlined and targeted predictive tools in a clinical environment, focusing on the most relevant factors for CCU readmission risk.

This stepwise selection approach was preferred over methods like principal component analysis (PCA)[21] and regression-based selection due to its unique benefits. Unlike PCA, which transforms the feature space into a set of linearly uncorrelated components and might obscure the interpretability of individual features, forward stepwise selection maintains the original features, thus preserving their interpretability.[22] Furthermore, while regression-based techniques might introduce all available features into the model simultaneously, potentially leading to overfitting, the stepwise method selectively incorporates only those features that provide meaningful improvements. The stepwise approach, therefore, achieves a balance between model complexity and predictive power, ensuring the inclusion of features that are both interpretable and relevant to the model's predictive capacity, an equilibrium that is less readily achieved with PCA or regression-based feature inclusion methods.

To make the model more interpretable in clinical settings, we adopted SHAP values for model explanation. SHAP values help to illustrate not only the importance of features but also how these features affect the model's predictions.[20,21] For instance, a longer length of CCU stay, ventilator use, and the presence of MI and atrial fibrillation are top important features associated with higher readmission risk, as indicated by positive SHAP values. By using the model interpretation process, clinicians can more easily understand how specific clinical variables contribute to the predictions.

The strategic adjustment of the decision threshold significantly impacts our model's ability to predict CCU readmissions. Setting a threshold of 0.025 enhances our model's sensitivity without significantly compromising specificity, enabling a more comprehensive detection of at-risk patients, as indicated in Table 3. Conversely, with the threshold at 0.125, our model achieves a high specificity of 0.90 in the hold-out test set, ensuring accurate identification of nonreadmissions while accepting a reduced sensitivity of 0.534. A similar pattern of sensitivity and specificity was observed across three other external validation hospitals. These adjustments highlight the delicate balance between capturing true positives and minimizing false alerts, allowing clinicians to tailor the model's use to the clinical scenario. In the hold-out test set, patients with model predictions greater than 0.125 demonstrated a 7.5-fold increased risk of CCU readmission compared to those with predictions lower than 0.125. In contrast, patients with prediction scores lower than 0.025 had a CCU readmission rate of only 0.5%. The decision on the optimal threshold must be carefully calibrated to the clinical setting, weighing the urgency of detecting readmissions against the consequences of false alarms. A lower threshold, such as 0.025, is prudent in high-stakes environments where a missed readmission could lead to serious consequences, while a higher threshold, like 0.125, may be more suitable in contexts where the costs of false positives are of greater concern.

This study has several limitations. First, due to its retrospective nature, data quality issues were present at both the development and external validation sites. Although the external validation hospitals and the primary hospital utilize the same EHR system, which facilitates consistent documentation of included features, assessing the model's performance in hospitals with different EHR systems poses a challenge. Second, in this study, we excluded patients who died or were transferred to other hospitals, which might introduce survival bias. This could potentially skew the data towards cases with less severity and affect the model's ability to generalize to sicker patients. Third, the calculation and justification of the sample size were not performed; this should be considered when interpreting the results. Furthermore, the clinical application of this model encounters limitations. The subset of features selected from the GB model primarily consisted of nonmodifiable characteristics from a clinician's perspective. Modifiable features that were shown to be associated with patient's outcome, like vital signs and medications were limited.[31,32] And despite the model's relatively PPVs, its ability to effectively rule out CCU readmissions remains a valuable asset. It can provide healthcare practitioners with a means to more efficiently allocate resources and enhance the care for patients less likely to experience readmission. Moreover, the implementation of any model into clinical practice necessitates ongoing surveillance and potential recalibration to suit the local context.

Conclusion

This study demonstrates the potential of ML-driven methods in predicting imminent CCU readmission during the initial hospital stay, leveraging comprehensive datasets from four prominent hospitals. Employing ML models, the research effectively utilized EHR features to optimize prediction performance. The GB model emerged superior, exhibiting remarkable performance with an AUC of 0.879 in the hold-out test set and external validations confirmed the model's robustness across different patient demographics and settings. Overall, this work reinforces ML's promise in enhancing patient care, though continuous refinement and validation are essential for clinical implementation.

Acknowledgements: We appreciated the Biostatistics Center, Kaohsiung Chang Gung Memorial Hospital for statistics consultation.

Contributorship: FFFY and IMC designed the model and the computational framework and analyzed the data. KHW and CYC carried out the implementation. IMC performed the calculations. FFFY, WCL, HCC, and CIC wrote the manuscript with input from all authors. TYC conceived the study and was in charge of overall direction and planning.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Ethical approval was obtained from the Institutional Review Board of Chang Gung Medical Foundation. The approval reference number is 202201541B0.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was funded by Chang Gung Medical Foundation (grant number: CMRPG8N0271).

Guarantor: TYC

Data availability statement: The data underlying this article will be shared on reasonable request to the corresponding author.

ORCID iD: I-Min Chiu https://orcid.org/0000-0001-6876-8217

Supplemental material: Supplemental material for this article is available online.

Reference

1. Bastero-Miñón P, Russell JL and Humpl T. Frequency, characteristics, and outcomes of pediatric patients readmitted to the cardiac critical care unit. *Intensive Care Med* 2012; 38: 1352–1357.
2. Chiang C-Y, Cheng F-J, Huang Y-S, et al. ICU admission following an unscheduled return visit to the pediatric emergency department within 72 hours. *BMC Pediatr* 2019; 19: 1–7.
3. Chen T-Y, Tseng C-H, Wu P-J, et al. Risk stratification model for predicting coronary care unit readmission. *Front Cardiovasc Med* 2022; 9: 825181. doi:10.3389/fcvm.2022.825181
4. Hassanien Ahmed A, Mohamed Mahany M, Abdelhamid Naguib T, et al. Nurses and physicians perception of causes and risk ractors contributing patient readmission to the coronary care unit. *Assiut Sci Nurs J* 2020; 8: 12–20.
5. Stewart S and Voss DW. A study of unplanned readmissions to a coronary care unit. *Heart Lung* 1997; 26: 196–203.
6. Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes* 2016; 9: 629–640.
7. Min X, Yu B and Wang F. Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: a case study on COPD. *Sci Rep* 2019; 9: 2362.
8. Hammoudeh A, Al-Naymat G, Ghannam I, et al. Predicting hospital readmission among diabetics using deep learning. *Procedia Comput Sci* 2018; 141: 484–489.
9. Huberts LCE, Li S, Blake V, et al. Predictive analytics for cardiovascular patient readmission and mortality: an explainable approach. *Comput Biol Med* 2024; 174: 108321.
10. Rojas JC, Carey KA, Edelson DP, et al. Predicting intensive care unit readmission with machine learning using electronic health record data. *Ann Am Thorac Soc* 2018; 15: 846–853.
11. Cheng C-Y, Chiu I-M, Zeng W-H, et al. Machine learning models for survival and neurological outcome prediction of out-of-hospital cardiac arrest patients. *BioMed Res Int* 2021; 2021(1): 9590131. doi:10.1155/2021/9590131
12. Kulenovic A and Lagumdzija-Kulenovic A. Using logistic regression to predict long COVID conditions in chronic patients. *Stud Health Technol Inform* 2022; 295: 265–268.
13. Chiu I-M, Cheng C-Y, Zeng W-H, et al. Using machine learning to predict invasive bacterial infections in young febrile infants visiting the emergency department. *J Clin Med* 2021; 10: 1875.
14. Montazeri M, Montazeri M, Montazeri M, et al. Machine learning models in breast cancer survival prediction. *Technol Health Care* 2016; 24: 31–42.
15. Ooka T, Johno H, Nakamoto K, et al. Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: large-scale health check-up data in Japan. *BMJ Nutr Prev Health* 2021; 4: 140–148.
16. Chiu I-M, Zeng W-H, Cheng C-Y, et al. Using a multiclass machine learning model to predict the outcome of acute ischemic stroke requiring reperfusion therapy. *Diagnostics* 2021; 11: 80.
17. Suthaharan S and Suthaharan S. Support vector machine. In: *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*. springer, 2016, pp.207–235.
18. Tsai C-M, Lin CHR, Zhang H, et al. Using machine learning to predict bacteremia in febrile children presented to the emergency department. *Diagnostics* 2020; 10: 307.
19. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001; 29(5): 1189–1232.
20. Yang Y, Wang Y-M, Lin C-HR, et al. Explainable deep learning model to predict invasive bacterial infection in febrile young infants: a retrospective study. *Int J Med Inf* 2023; 172: 105007.
21. Hsu T-Y, Cheng C-Y, Chiu I-M, et al. Explainable deep learning model for predicting serious adverse events in hospitalized geriatric patients within 72 hours. *Clin Interv Aging* 2024; 19(June): 1051–1063. doi:10.2147/CIA.S460562
22. Liu H and Motoda H. *Feature selection for knowledge discovery and data mining*, Springer Science & Business Media, 2012, Vol. 454.
23. Kino S, Hsu Y-T, Shiba K, et al. A scoping review on the use of machine learning in research on social determinants of health: trends and research prospects. *SSM Popul Health* 2021; 15: 100836.
24. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; 366: 447–453.
25. Singer DE, Mulley AG, Thibault GE, et al. Unexpected readmissions to the coronary-care unit during recovery from acute myocardial infarction. *N Engl J Med* 1981; 304: 625–629.
26. Kessler S, et al. Predicting readmission to the cardiovascular intensive care unit using recurrent neural networks. *Digit Health* 2023; 9: 20552076221149529.
27. Salimi M, Bastani P, Nasiri M, et al. Predicting readmission of cardiovascular patients admitted to the CCU using data mining techniques. *Open Cardiovasc Med J* 2023; 17: e187419242305170.
28. Wang Y-M, Chiu I-M, Chuang Y-P, et al. RAPID-ED: a predictive model for risk assessment of patient's early in-hospital deterioration from emergency department. *Resusc Plus* 2024; 17: 100570.
29. Franchi C, Nobili A, Mari D, et al. Risk factors for hospital readmission of elderly patients. *Eur J Intern Med* 2013; 24: 45–51.
30. Chiang C-Y, Chen Y-L, Lin Y-R, et al. Characteristics of febrile children admitted to the ICU following an unscheduled ED revisit within 72 h, a case–control study. *Front Pediatr* 2020; 8: 411.
31. Huang Y-S, Chiu I-M, Tsai M-T, et al. Delta shock index during emergency department stay is associated with in hospital mortality in critically ill patients. *Front Med (Lausanne)* 2021; 8: 648375.
32. Huang Y-H, Yan J-H, Kuo K-C, et al. Early antibiotics use in young infants with invasive bacterial infection visiting emergency department, a single medical center's experience. *Pediatr Neonatol* 2020; 61: 155–159.