Shan *et al. International Journal of Emergency Medicine* (2025) 18:154
https://doi.org/10.1186/s12245-025-00947-8

International Journal of
Emergency Medicine

RESEARCH **Open Access**

The association between the severity of out-of-hospital cardiac arrest and the effectiveness of target temperature management: a retrospective study based on prediction models

Zhenyu Shan[1], Rui Shao[1], Xingsheng Wang[1], Guyu Zhang[1], Luying Zhang[1], Chenchen Hang[1], Le An[1], Jingfei Yu[1] and Ziren Tang[1*]

Abstract

Aim This study aimed to develop prediction models and conduct risk stratifications for patients with out-of-hospital cardiac arrest (OHCA) to identify patients who could benefit from targeted temperature management (TTM) at 33°C.

Methods A retrospective analysis was carried out on 368 patients and the primary outcome was the neurological outcome at discharge evaluated by the Cerebral Performance Categories (CPC) scale. Six variables were utilized to construct prediction models via six methodologies, and the Chi-square test or Fisher's exact test was used to analyze the efficacy of TTM at 33℃ under diverse risk stratifications.

Results A total of 264 eligible patients were divided into the development cohort and test set. The identified predictors comprised bystander cardiopulmonary resuscitation (CPR), pupillary light reflex, Acute Physiology and Chronic Health Evaluation II (APACHE II) score, lactate, serum calcium (Ca^{2+}), and base excess (BE). The AUC of different prediction models in the test set ranged from 0.7592 to 0.9304. Patients with a predicted probability of 80-100%, 75-100%, and 67-100% in the Random Forest model, and 40-60% in the K-Nearest Neighbors model, can benefit from 33℃ TTM (OR [95% CI]: 3.21[1.44–7.19], 2.73[1.25–5.97], 2.18[1.09–4.36], 6.42[1.09–37.73], respectively). Among patients who had successfully undergone TTM at 33 °C, there was a higher prevalence of patients classified as CPC 3 and CPC 4 and a lower incidence of those classified as CPC 5 (OR [95% CI]: 3.90[1.12–12.58], 2.29[1.24–4.26], 0.31[0.19–0.51], respectively).

Conclusion Prediction models developed from early variables can predict the neurological prognosis of OHCA, and the efficacy of 33℃ TTM may be related to severity.

---

*Correspondence:
Ziren Tang
tangziren1970@163.com

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

Keywords Out-of-hospital cardiac arrest, Return of spontaneous circulation, Prediction models, Severity, Target temperature management

Introduction

Out-of-hospital cardiac arrest (OHCA) constitutes a major global health concern, with the annual incidence varying from 30 to 97 per 100,000 population [1]. Although the rate of return of spontaneous circulation (ROSC) is around 30% [2], the proportion of favorable neurological outcomes is merely 2.8–18.2% [1]. In China, less than 1% of patients are discharged with good neurological function [3]. This severe situation and striking discrepancy underscore the urgent need to optimize post-resuscitation strategies for OHCA.

Targeted temperature management (TTM) has long been regarded as a therapeutic protocol to mitigate cerebral ischemia-reperfusion injury after OHCA [4]. Early animal studies [5] demonstrated that hypothermia reduces excitotoxicity, oxidative stress, and mitochondrial dysfunction, laying the theoretical foundation for its clinical application. However, translating these findings into practice has been complicated by conflicting evidence regarding optimal target temperatures.

Multicenter randomized controlled trials [6, 7] have shown no significant difference in the prognosis of OHCA patients under different TTM. The TTM1 Trial compared comatose OHCA survivors treated with target temperatures of 33°C and 36°C, and found no significant difference in neurological outcomes at 180 days. This was further supported by the TTM2 Trial, which demonstrated that normothermia was non-inferior to hypothermia. However, several large retrospective studies [8, 9] have indicated that the efficacy of TTM is influenced by the severity of OHCA. Specifically, TTM at 32°C-36°C was significantly associated with a better outcome in OHCA patients with low and high severity, and TTM at 36°C was associated with better survival among patients with the mild to moderate severe post-cardiac arrest syndrome (PCAS). Despite these insights, current clinical practice lacks tools to identify which patients would benefit most from 33°C versus 36°C. Existing prediction models for OHCA outcomes primarily focus on risk stratification but do not integrate TTM responsiveness [10, 11].

The objective of this study was to develop and validate prediction models for adult patients with OHCA, and conduct risk stratifications, aiming to identify those who could benefit from TTM at 33°C. Specifically, these models will also predict neurological outcomes immediately following the achievement of ROSC, thereby providing evidence for the application of TTM under different risk stratifications.

Methods

Study population

During the period from August 1, 2016, to September 30, 2024, a total of 368 cardiac arrest (CA) patients were admitted to the Emergency Intensive Care Unit (EICU). The inclusion criteria were patients confirmed by emergency medical services personnel as having suffered OHCA. The exclusion criteria were as follows: traumatic brain injury (such as subarachnoid hemorrhage, cerebral hemorrhage, etc.), pregnancy, age under 18 years, use of extracorporeal cardiopulmonary resuscitation (ECPR), failure to achieve ROSC, and missing critical pre-hospital data. Based on these criteria, 264 OHCA patients were included. The treatment and care after cardiac arrest are guided by the American Heart Association Guidelines [12], implemented in accordance with the treatment recommendations of the International Liaison Committee on Resuscitation [13] and practical clinical considerations, and updated in line with revisions to the guidelines and treatment recommendations. The study was approved by the Institutional Review Board of Beijing Chao-yang Hospital.

Data collection, definitions, and missing values

Retrospective data collection was performed for each patient under the Utstein style [14]. The collected data encompassed the following variables: age, gender, cause of cardiac arrest (CA), witnessed status, bystander cardiopulmonary resuscitation (CPR), initial heart rhythm, time from collapse/discovery to ROSC (TROSC), coronary angiogram/percutaneous coronary intervention (CAG/PCI), TTM, Acute Physiology and Chronic Health Evaluation II (APACHE II) [15] score, sequential organ failure assessment (SOFA) [16] score, vital sign, laboratory examination, etc. The etiology was presumed to be cardiac unless it was definitively known or highly likely to have been caused by trauma, submersion, drug overdose, asphyxia, or any other non-cardiac cause. Successful implementation of TTM at 33 °C was defined by meeting the following criteria: immediate initiation of TTM upon admission to EICU, a target temperature set at 33 °C, absence of hemodynamic instability or death during the TTM process, and successful completion of the rewarming phase, with or without subsequent continuation of temperature control. Patients who met these criteria for successful TTM at 33°C were categorized as the TTM Data group, and the remaining patients were designated as the non-TTM Data group.

We excluded in-hospital variables with more than 20% missing values. For missing values of other variables, we used Predictive Mean Matching (PMM) for imputation via R software (‘mice’ package) (Supplementary Information).

Outcomes

The primary outcome of this study was the neurological outcome at discharge, which was assessed by a single assessor (Z.S.) using the Cerebral Performance Categories (CPC) scale (CPC 1, good cerebral performance; CPC 2, moderate cerebral disability; CPC 3, severe cerebral disability; CPC 4, coma or vegetative state; CPC 5, death/brain death). CPC 1–2 denoted a good neurological outcome, and CPC 3–5 denoted a poor neurological outcome.

Statistical analysis

Categorical variables were described as frequencies with percentiles (%). Continuous variables were summarized as mean with standard deviations (SD) (if normal distribution) or median with interquartile ranges [IQRs] (if not normal distribution). Comparisons used the Chi-square test or Fisher’s exact test for categorical variables and Student’s t-test or the Mann-Whitney U test for continuous variables.

Patients prior to May 1, 2023, were designated as the development cohort, and the remaining patients were regarded as the test set. The development cohort was randomly partitioned into the train set and validation set in a ratio of 7:3 (Fig. 1). In the development cohort, univariate logistic regression and the Least Absolute Shrinkage and Selection Operator (LASSO) were implemented for the selection of significant variables [17]. In the LASSO regression, a regularization term is added to the loss function, which is the sum of the absolute values of the model coefficients multiplied by lambda. We perform 10-fold cross-validation to estimate the value of lambda and use lambda.1se for the sake of model simplicity using the “glmnet” package (Supplementary Information).

To evaluate the performance and predictive value of different models, six methods were applied in the model construction process, including Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), eXtreme Gradient Boosting (XGBoost), K-Nearest Neighbors (k-NN), and Support Vector Machine (SVM). The modeling process was carried out in the train set, and hyperparameter

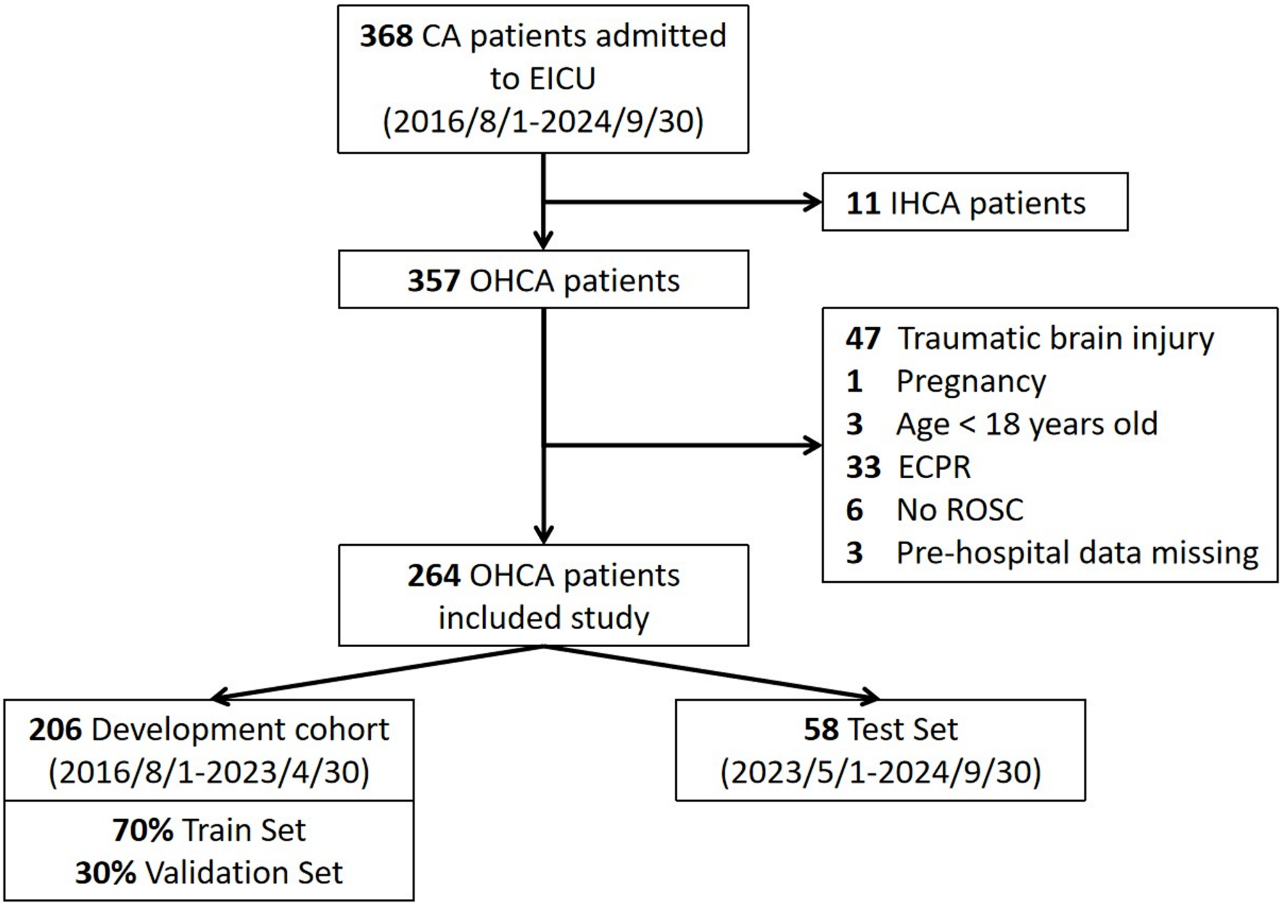

Fig. 1 Flowchart of study

tuning was conducted in the validation set to attain optimal C-statistics (Supplementary Information). The ROC curves were plotted, and the area under the curve (AUC) of different models was calculated. Additionally, 400 times of 5-fold cross-validation were carried out in the development cohort, and the test set was fitted into 2000 repetitions of the modeling process. We also computed sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and F1-score based on the threshold value corresponding to the best Youden index and a predicted probability greater than 0.5, respectively. The F1-score was calculated based on precision and recall. DeLong's test was performed to assess whether there were differences in AUC among different models and datasets. The calibration curve and decision curve of the model were also plotted, and the Hosmer-Lemeshow goodness of fit test, Chi-Square goodness of fit test, and Monte Carlo Simulation were utilized to determine whether the model was well calibrated.

In this study, the patients were respectively divided into 2 groups, 3 groups, 4 groups, 5 groups, and 10 groups according to the predicted probability to ascertain whether patients could benefit from different TTM settings (For example, being divided into 4 groups means that the prediction probabilities for each group are 0–25%, 25-50%, 50-75%, and 75-100%, respectively). The effectiveness of TTM on CPC score was analyzed using the Chi-square test or Fisher's exact test.

All the statistical analyses were performed using the R software version 4.3.1 (R Development Core Team, 2023) with RStudio version 2023.12.1. All tests were two-tailed, and a p-value of less than 0.05 was considered statistically significant.

Results

Study population and characteristics

A total of 264 eligible patients were divided into the development cohort ($n=206$) and the test set ($n=58$) (Fig. 1). Between the development cohort and test set, the following variables were different: successful TTM at 33℃ (n, 60(29.1%) vs. 16(27.6%)), APACHE II (28[24.0–32.0] vs. 31[27–36]), and Lactate (mmol/L, 8.1[4.1–13.3] vs. 10.6[5.5–14.3]) (All $P<0.05$). In the development cohort, 160(77.7%) patients were discharged with CPC 3–5, in comparison with 47(81.1%) patients in the test set (Table 1 and Table S1).

Selection of predictors

Univariate logistic regression analysis

The variables that exhibited statistical significance were identified as follows: age, cardiac etiology, bystander CPR, initial rhythm, "no-flow" time, "low-flow" time, TROSC, defibrillation, adrenaline, CAG/PCI, APACHE II, SOFA, pupillary light reflex, respiratory rate, heart rate, withdrawal of invasive therapy, white blood cell, lactate, pH, $PaCO_2$, base excess (BE), total protein, albumin, high-density lipoprotein, low-density lipoprotein, aspartate aminotransferase, lactate dehydrogenase, blood urea nitrogen, creatinine, uric acid, serum calcium (Ca^{2+}), phosphatase, anion gap, glucose, prothrombin time, prothrombin time activity, international normalized ratio, and D-Dimer (Table 2 and S1).

LASSO regression

Among the 38 variables enumerated above, we excluded CAG/PCI, respiratory rate, heart rate, and withdrawal of invasive therapy. This exclusion was based on considerations related to potential differences in treatment regimens, the use of vasoactive drugs, and ventilator management. Under the regularization process of LASSO and lambda.1se criterion, six predictors were identified, namely bystander CPR, pupillary light reflex, APACHE II, lactate, Ca^{2+}, and BE, presented in the order of variable importance as determined by the LASSO Regression (Supplementary Information).

Development and performance of the prediction models

The six variables above were utilized to construct the prediction models. The modeling techniques encompassed LR, RF, DT, XGBoost, k-NN, and SVM (Supplementary Information).

ROC curves and other performance measures of prediction models

The AUC of the six models ranged from 0.9356 to 0.9553 in the train set, from 0.9215 to 0.9575 in the validation set, and from 0.7592 to 0.9304 in the test set (Fig. 2). In the test set, significant differences were observed between the ROC curves of the DT model and those of the LR, k-NN, and SVM models ($P=0.029$, 0.026, 0.024, respectively) (Table S2). The AUC values of cross-validation varied from 0.9011 to 0.9377, while the AUC values of the test set ranged from 0.6846 to 0.9212 (Figure S2). Sensitivity, specificity, PPV, NPV, accuracy, and F1-score are described in Fig. 3 and Table S3. It is noteworthy that the specificity of the RF model was the lowest across all data sets.

Calibration curves and decision curves of prediction models

The calibration curves of these prediction models were plotted (Figures S3 and S4). The goodness of fit test indicated that only in the RF model, the calibration was poor when the prediction probability was less than 40% ($P<0.05$) (Table S4). Figure S5 presents the decision curves for the six models (Supplement Information).

Table 1 Baseline characteristics in the development cohort and test set

Variable	**Development Cohort**			***P*-value***	**Test Set (*n* = 58)**	***P*-value†**
	All Data (*n* = 206)	**CPC 1–2 (*n* = 46)**	**CPC 3–5 (*n* = 160)**			
Age, year	64[49.3, 74.0]	57.5[40.3, 70.3]	66[51, 74.5]	0.0194	64[53, 71.8]	0.9488
Male, *n* (%)	134(65)	26(56.5)	108(67.5)	0.2298	34(58.6)	0.4566
Cardiac etiology, *n* (%)	121(58.7)	35(76.1)	86(53.7)	0.0110	29(50.0)	0.2999
Bystander CPR, *n* (%)	144(69.9)	44(95.7)	100(62.5)	< 0.0001	39(67.2)	0.8204
Initial rhythm, *n* (%)				0.0023		0.2881
VF/VT	67(32.5)	24(52.2)	43(26.9)		14(24.1)	
PEA/Asystole	139(67.5)	22(47.8)	117(73.1)		44(75.9)	
TROSC, min	25[15.0, 48.8]	15[9.3, 25.0]	30[19.8, 53.8]	< 0.0001	30[17.8, 55.3]	0.2732
Defibrillation, *n* (%)	75(36.4)	26(56.5)	49(30.6)	0.0023	21(36.2)	1
CAG/PCI, *n* (%)	33(16)	15(32.6)	18(11.2)	0.0011	8(13.8)	0.8350
TTM, *n* (%)	98(47.6)	25(54.3)	73(45.6)	0.3808	16(27.6)	0.0103
APACHE II	28[24.0, 32.0]	21[17.0, 24.8]	30[26.0, 33.0]	< 0.0001	31[27, 36]	0.0014
SOFA	10.6 ± 4.1	7.5 ± 3.9	11.5 ± 3.7	< 0.0001	11.3 ± 3.3	0.1857
Pupillary light reflex, *n* (%)				< 0.0001		0.8301
Sensitive	59(28.6)	28(60.9)	31(19.4)		15(25.9)	
Sluggish	74(35.9)	15(32.6)	59(36.9)		20(34.5)	
Absent	73(35.4)	3(6.5)	70(43.8)		23(39.7)	
CPC, *n* (%)						0.7129
1	31(15)	31(67.4)			6(10.3)	
2	15(7.3)	15(32.6)			5(8.6)	
3	13(6.3)		13(8.1)		2(3.4)	
4	42(20.4)		42(26.3)		10(17.4)	
5	105(51)		105(65.6)		35(60.3)	
Length of stay, hour	289[87.0, 541.8]	416[254.8, 588.8]	258[49.0, 520.5]	0.0023	234[107.0, 382.2]	0.2110
Lactate, mmol/L	8.1[4.1, 13.3]	2.6[1.3, 5.2]	9.6[5.8, 14.0]	< 0.0001	10.6[5.5, 14.3]	0.0360
pH	7.24[7.09, 7.35]	7.35[7.31, 7.40]	7.18[7.03, 7.30]	< 0.0001	7.23[7.07, 7.33]	0.4727
BE, mmol/L	−7.7[−12.8, −3.1]	−3.1[−5.7, −1.6]	−9.3[−15.0, −5.0]	< 0.0001	−8.4[−13.1, −3.2]	0.4670
Ca^{2+}, mmol/L	1.98[1.84, 2.10]	2.12[2.00, 2.19]	1.93[1.81, 2.06]	< 0.0001	2.02[1.93, 2.14]	0.0881

Data are shown as median [IQR] or mean ± SD unless otherwise indicated

CPR cardiopulmonary resuscitation, *VF/VT* ventricular tachycardia/fibrillation, *PEA* pulseless electrical activity, *TROSC* time to return of spontaneous circulation, *CAG* coronary angiogram, *PCI* percutaneous coronary intervention, *BE* base excess, *CPC* Cerebral Performance Categories

*CPC 1–2 vs. CPC 3–5 in Development Cohort

†Development Cohort vs. Test Set

Effectiveness of TTM

Among the overall patients ($n = 264$), there was no significant difference in the neurological functional prognosis between the patients in the TTM Data and those in the non-TTM Data (OR [95% CI]: 1.63[0.90–2.93], $P = 0.1400$). Given that the efficacy of TTM might be influenced by the severity of OHCA, this study analyzed the association under different prediction models.

Predictive performance of prediction models in TTM and non-TTM data

TTM and non-TTM Data were incorporated into the aforementioned prediction models, and the ROC and calibration curves are presented in Figure S5 and Figure S6. The results indicate that the AUC for TTM data ranges from 0.902 to 0.9325, while the AUC for non-TTM data spans from 0.88 to 0.9654. DeLong's test demonstrated that there was no significant difference between the two (all $p > 0.05$) (Table S2).

Effectiveness of TTM under different risk stratifications

According to the predicted probability, the patients were divided into 2 groups, 3 groups, 4 groups, 5 groups, and 10 groups. The number of patients in each group and the specific results are presented in Table S5 and Figure S8. The analysis revealed that in the RF model when the predicted probabilities were 80-100% (5 Groups), 75-100% (4 Groups), and 67-100% (3 Groups), patients were likely to benefit from TTM at 33℃ (OR [95% CI]: 3.21[1.44–7.19], 2.73[1.25–5.97], 2.18[1.09–4.36], $P = 0.0062$, 0.0166, 0.0333, respectively). In the k-NN model, when the predicted probability was within the range of 40-60% (5 Groups), patients may benefit more from TTM at 33℃ (OR [95% CI]: 6.42[1.09–37.73], $P = 0.0486$), as illustrated

Table 2 Univariate logistic analysis and LASSO regression coefficients

Variable	**Univariate logistic analysis**		**LASSO regression**
	OR (95% CI)	***P*-value**	**coefficient**
Age	1.025(1.005–1.045)	0.0126	0
Female	0.626(0.321–1.232)	0.1710	0
Non-Cardiac etiology	2.738(1.335–6.001)	0.0081	0
Non-Bystander CPR	13.20(3.877–82.71)	0.0005	0.2833
Non-shockable rhythm	2.968(1.512–5.876)	0.0016	0
TROSC	1.045(1.024–1.071)	0.0002	0
APACHE II	1.351(1.239–1.502)	< 0.0001	0.1194
SOFA	1.347(1.214–1.515)	< 0.0001	0
Pupillary light reflex (Sensitive)			0.2361
Sluggish	3.553(1.679–7.775)	0.0011	
Absent	21.08(6.826–92.85)	< 0.0001	
Lactate	1.423(1.276–1.621)	< 0.0001	0.1026
BE	0.872(0.818–0.924)	< 0.0001	−0.0005
Ca^{2+}	0.028(0.004–0.149)	< 0.0001	−0.0939

CPR cardiopulmonary resuscitation, *TROSC* time to return of spontaneous circulation, *BE* base excess

in Fig. 4. For other models and stratifications, such a difference was not observed.

Correlation between TTM and CPC scores

This study examined the distribution of CPC scores of TTM Data and non-TTM Data. The results indicated a significant difference in the overall CPC scores of patients ($P = 0.0001$). A comparison of CPC scores across each category revealed that there was no disparity between CPC 1 and CPC 2 in the overall patient population. Meanwhile, the incidences of CPC 3 and CPC 4 were higher in TTM Data compared to non-TTM Data (OR [95% CI]: 3.90[1.12–12.58], 2.29[1.24–4.26], $P = 0.0308$, 0.0120, respectively), and the reverse was the case for CPC 5 (OR [95% CI]: 0.31[0.19–0.51], $P < 0.0001$) (Fig. 5).

Discussion

In this study, we developed the prediction models based on six variables derived from early post-resuscitation data to predict the neurological outcomes at discharge. Additionally, we stratified the disease risk

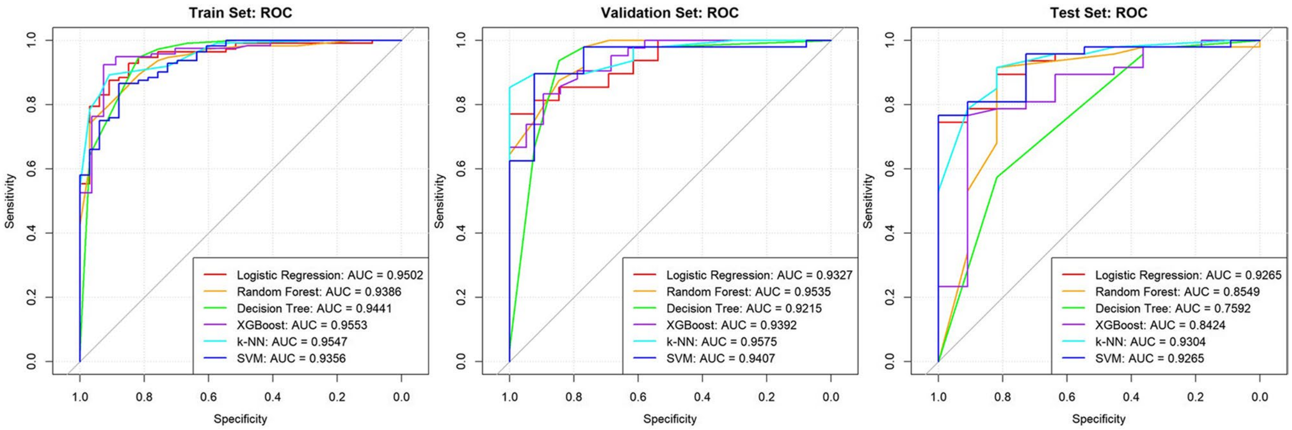

Fig. 2 ROC curves and AUCs of the train set, validation set, and test set. AUC Area Under Curve, XGBoost eXtreme Gradient Boosting, k-NN K-Nearest Neighbors, SVM Support Vector Machine

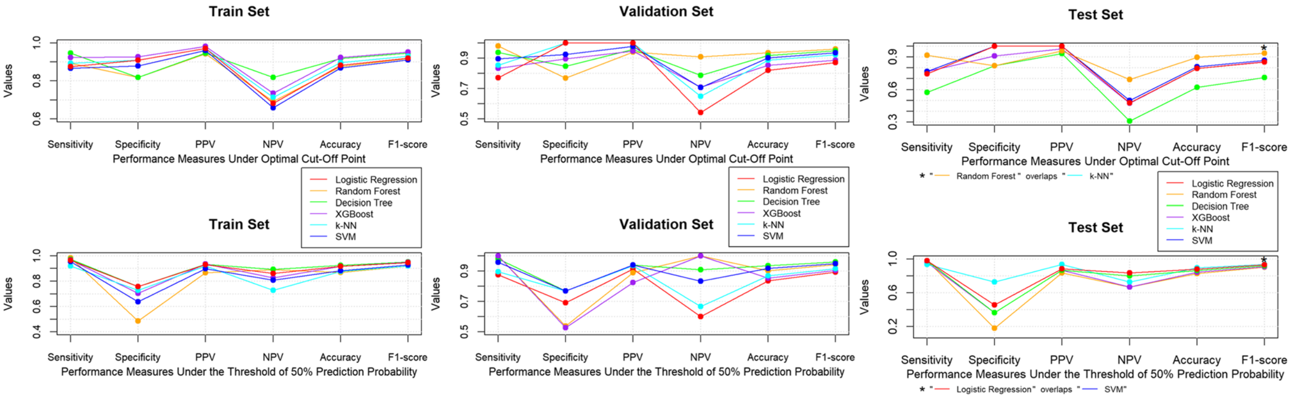

Fig. 3 Performance measures of different prediction models in the train set, validation set, and test set. * The two lines overlap. XGBoost eXtreme Gradient Boosting, k-NN K-Nearest Neighbors, SVM Support Vector Machine, PPV positive predictive value, NPV negative predictive value

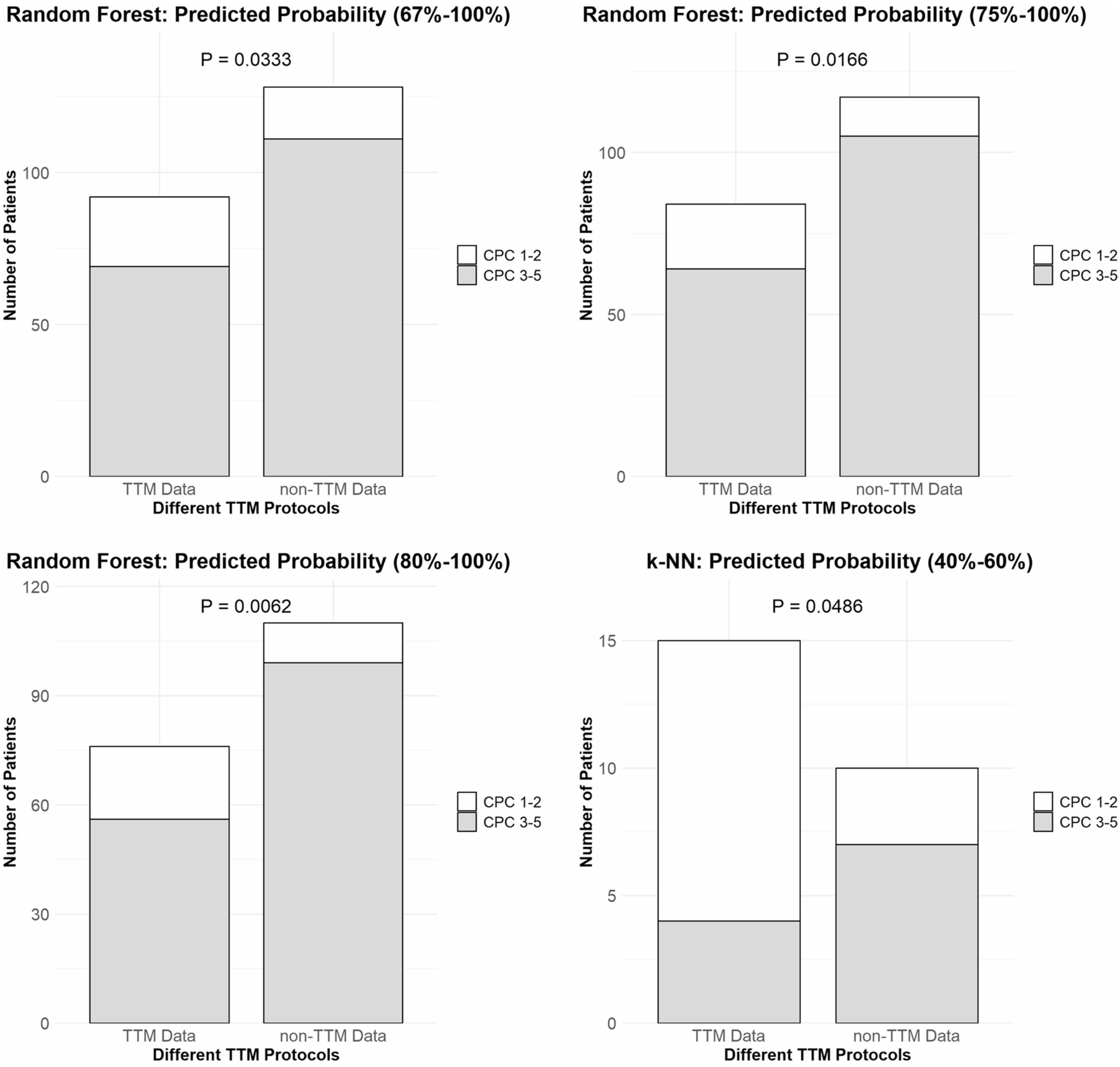

Fig. 4 Correlation between severity of PCAS and effectiveness of TTM

according to the predicted probability and identified the patients who could benefit from TTM at 33°C.

The predictors of this study encompassed bystander CPR, pupillary light reflex, APACHE II, lactate, Ca^{2+}, and BE. Bystander CPR reflects the immediacy of the patient receiving BLS. Early and effective implementation of CPR can enhance the prognosis of patients and has been widely incorporated into the development of prediction models [18–21]. The pupillary light reflex represents a straightforward neurological function assessment, which has been demonstrated to be correlated with the prognosis of OHCA patients [22, 23]. The APACHE II score serves as a comprehensive tool for assessing disease severity, consisting of three components: an acute physiological score, a chronic health score, and an age score. It is routinely calculated upon admission to the EICU. Previous studies have indicated that it can predict the prognosis of OHCA patients [24, 25]. Following OHCA, the body undergoes a severe disruption of acid-base metabolism. Lactate and BE can respectively reflect the tissue hypoxia level and metabolic acid-base imbalance, thereby facilitating the evaluation of disease severity [22]. In this study, we incorporated a relatively novel predictor, namely Ca^{2+}, which represents the concentration of total calcium in serum and plays a pivotal role in

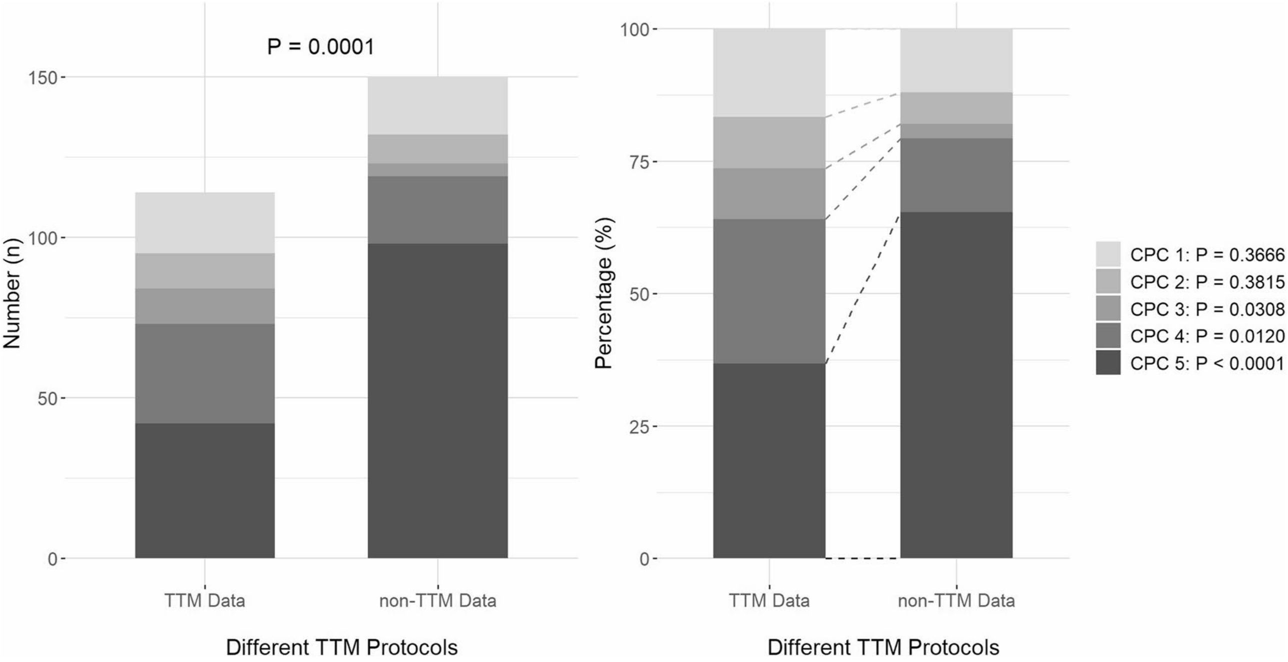

Fig. 5 CPC scores in different groups

safeguarding the stability of the myocardial membrane [26]. Previous studies have demonstrated that calcium is correlated with the prognosis of multiple diseases, with CA being among them [27–29]. Nevertheless, there is currently no evidence to suggest that routine calcium supplementation can enhance the outcomes of CA [30]. Further investigations are warranted to ascertain the predictive significance of Ca^{2+} in patients with OHCA. Utilizing these predictors as a foundation, we developed the prediction models.

In this study, six commonly utilized modeling strategies were simultaneously adopted. Through comparative analysis, it was found that AUC of most modeling methods in the validation set exceeded 0.84, with the exception of the DT model. In terms of model calibration, all models demonstrated satisfactory calibration, except when the prediction probability of the RF model was less than 40%, where it exhibited poor calibration. This, in turn, led to the poorest specificity for this model among all those considered. Consequently, we posit that these models possess the capability to predict unfavorable neurological outcomes. Subsequently, we incorporated TTM Data and non-TTM Data into the model respectively and conducted DeLong's test. The results of the test indicated that the prediction of disease severity was independent of whether TTM at 33°C was implemented.

Considering that the efficacy of TTM may be correlated with disease severity [8, 31], we stratified the severity according to the predicted probability, and the results showed that the RF and k-NN models can distinguish which patients can benefit from 33°C TTM. Although the RF model was calibrated incorrectly when the predicted probability was < 40%, the conclusions were drawn in patients with a predicted probability > 67% and thus may be reliable. However, this distinction needs to be verified by prospective studies. This suggests that in the clinical application of TTM, the severity of OHCA should be taken into consideration. The study also revealed that although TTM did not lead to a significant improvement in neurological outcomes, it did contribute to enhanced discharge survival.

The study also presented certain limitations. First, the analysis in this study was predicated on retrospective data, and consequently, missing values were inevitable. Patient resuscitation details, such as "no-flow" time, "low-flow" time, and TROSC, are frequently utilized in model construction [11, 19, 20, 22]. However, in this study, such information could only be retrieved from medical records and might not be entirely accurate. This could potentially induce selection bias and information bias, thereby influencing the selection of variables. Second, the number of patients incorporated into the modeling process of this study was relatively modest, with 34 candidate variables. This circumstance might impede the optimal selection of predictor variables. Moreover, despite external validation, the fact that it was a single-center study might circumscribe the generalizability of the models. Hence, larger databases and participation from other centers are requisite for validating this model. Thirdly, over

50% of the patients exhibited a predicted probability of poor neurological outcomes exceeding 90% (Figure S8). Such data imbalance might have predisposed the model to overfit and led to a paucity of low-risk individuals within the stratification of disease severity. This, in turn, could have undermined the robustness of the conclusions. For illustration, in the efficacy analysis of the k-NN model, the Fragility Index was calculated to be 1, indicating that changing just one outcome event would alter the statistical significance of the conclusion. Fourthly, the outcome metric of this study was the CPC score at discharge, precluding the evaluation of long-term outcomes. Finally, this study excluded patients with brain injury and ECPR, who accounted for 12.8% and 9.0% of all CA patients, respectively. The prognosis and hemodynamics of these two cohorts of patients deviate markedly from those of other patients [32]. Therefore, it is imperative to ascertain the efficacy of TTM in these particular patient groups.

Conclusion

Our study demonstrates that the prediction models constructed using predictors at admission are capable of predicting the neurological outcome of OHCA. Additionally, the effect of TTM at 33℃ may be associated with the severity of PCAS.

Abbreviations

APACHE II	Acute Physiology and Chronic Health Evaluation II
AUC	Area Under Curve
BE	Base Excess
CA	Cardiac Arrest
Ca^{2+}	Serum Calcium (Free Calcium Ion + Protein-Bound Calcium)
CAG	Coronary Angiogram
CI	Confidential Interval
CPC	Cerebral Performance Category
CPR	Cardiopulmonary Resuscitation
DT	Decision Tree
ECPR	Extracorporeal Cardiopulmonary Resuscitation
EICU	Emergency Intensive Care Unit
IHCA	In-Hospital Cardiac Arrest
k-NN	K-Nearest Neighbors
LR	Logistic Regression
OHCA	Out-of-Hospital Cardiac Arrest
OR	Odds Ratio
PCI	Percutaneous Coronary Intervention
$PaCO_2$	Partial Pressure of Carbon Dioxide in Artery
PEA	Pulseless Electrical Activity
pH	Potential of Hydrogen
RF	Random Forest
ROSC	Return of Spontaneous Circulation
SOFA	Sequential Organ Failure Assessment
SVM	Support Vector Machine
TROSC	Time to ROSC
TTM	Targeted Temperature Management
VF	Ventricular Fibrillation
VT	Ventricular Tachycardia
XGBoost	eXtreme Gradient Boosting

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12245-025-00947-8.

Supplementary Material 1: Table S1. Baseline characteristics of patients. Table S2. DeLong's test of different models and different data sets. Table S3. Performance measure of different models. Table S4. Hosmer-Lemeshow goodness of fit test and Chi-Square goodness of fit test. Table S5. Different groups of patients based on predicted probability. Figure S1. LASSO Regression. Left: A coefficient path plot was produced against the Log Lambda sequence. Right: Cross-validation plot for the penalty term. Six variables were selected according to lambda.1se. Figure S2. ROC curves and AUCs for cross-validation and test set. Figure S3. Calibration curves of the train set, validation set, and test set. Figure S4. Calibration curve of cross-validation and test set. Figure S5. Decision curves of training set, validation set, and test set. DCA Decision Curve Analysis, All interfere with all patients, None not interfere with anyone. Figure S6. ROC curve and AUCs of TTM Data and non-TTM Data. Data represents all patients, i.e., TTM Data + non-TTM Data. Figure S7. Calibration curves of TTM Data and non-TTM Data. Data represents all patients, i.e., TTM Data + non-TTM Data. Figure S8. Different groups of patients based on predicted probability. Data represents all patients, i.e., TTM Data + non-TTM Data. The pictures include 2 groups, 3 groups, 4 groups, 5 groups, and 10 groups. The solid line represents the line chart of observed probability. The dashed line represents the average observed probability. The box is the probability of occurrence for different predicted probabilities. The hollow point represents the observed probability that cannot be calculated, so the average of the preceding and following observed probability or the nearest observed probability is used.

Acknowledgements
We sincerely thank all authors of all primary clinical studies that were included in the references.

Authors' contributions
ZS: Writing – original draft, Writing - Review & Editing, Investigation, Methodology, Software, Conceptualization. RS: Writing - Review & Editing, Investigation. XW: Writing - Review & Editing, Investigation. GZ: Software, Investigation, Resources, Validation. LZ: Writing - Review & Editing, Investigation. CH: Writing - Review & Editing, Supervision. LA: Resources, Supervision. JY: Data Curation, Investigation. ZT: Writing - Review & Editing, Conceptualization, Supervision, Project administration, Funding acquisition. All authors actively participated in the development of the article and gave their approval for the final version to be submitted.

Funding
This research was funded by the following sources: High-Level Public Health Technical Talent Building Program (Discipline Leader-01-01); Beijing Hospitals Authority's Ascent Plan (DFL20240302).

Data availability
To protect the privacy of the research participants and in accordance with the requirements of the ethics committee, the original data of the participants were not provided in this study. The data analysed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the Hospital Institutional Review Board and the Ethics Committee of Beijing Chaoyang Hospital. The requirement for informed consent was waived because of the retrospective design.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Emergency Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China

Received: 13 May 2025 / Accepted: 3 July 2025

References

1. Kiguchi T, Okubo M, Nishiyama C, et al. Out-of-hospital cardiac arrest across the world: first report from the international liaison committee on resuscitation (ILCOR). Resuscitation. 2020;152:39–49.
2. Yan S, Gan Y, Jiang N, et al. The global survival rate among adult out-of-hospital cardiac arrest patients who received cardiopulmonary resuscitation: a systematic review and meta-analysis. Crit Care. 2020;24:61.
3. Zheng J, Lv C, Zheng W, et al. Incidence, process of care, and outcomes of out-of-hospital cardiac arrest in china: a prospective study of the BASIC-OHCA registry. Lancet Public Health. 2023;8:e923–32.
4. Polderman KH. Mechanisms of action, physiological effects, and complications of hypothermia. Crit Care Med. 2009;37:S186-202.
5. Arrich J, Herkner H, Müllner D, et al. Targeted temperature management after cardiac arrest. A systematic review and meta-analysis of animal studies. Resuscitation. 2021;162:47–55.
6. Dankiewicz J, Cronberg T, Lilja G, et al. Hypothermia versus normothermia after Out-of-Hospital cardiac arrest. N Engl J Med. 2021;384:2283–94.
7. Nielsen N, Wetterslev J, Cronberg T, et al. Targeted temperature management at 33°C versus 36°C after cardiac arrest. N Engl J Med. 2013;369:2197–206.
8. Callaway CW, Coppler PJ, Faro J, et al. Association of initial illness severity and outcomes after cardiac arrest with targeted temperature management at 36°C or 33°C. JAMA Netw Open. 2020;3:e208215.
9. Lascarrou JB, Dumas F, Bougouin W, et al. Differential effect of targeted temperature management between 32°C and 36°C following cardiac arrest according to initial severity of illness: insights from two international data sets. Chest. 2023;163:1120–9.
10. Carrick RT, Park JG, McGinnes HL, et al. Clinical predictive models of sudden cardiac arrest: A survey of the current science and analysis of model performances. J Am Heart Assoc. 2020;9:e017625.
11. Maupain C, Bougouin W, Lamhaut L, et al. The CAHP (Cardiac arrest hospital Prognosis) score: a tool for risk stratification after out-of-hospital cardiac arrest. Eur Heart J. 2016;37:3222–8.
12. Neumar RW, Shuster M, Callaway CW, et al. Part 1: Executive Summary: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation. 2015;132:S315-67.
13. Greif R, Bray JE, Djärv T, et al. 2024 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science With Treatment Recommendations: Summary From the Basic Life Support; Advanced Life Support; Pediatric Life Support; Neonatal Life Support; Education, Implementation, and Teams; and First Aid Task Forces. Circulation. 2024;150:e580-e687.
14. Perkins GD, Jacobs IG, Nadkarni VM, et al. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update of the Utstein Resuscitation Registry Templates for Out-of-Hospital Cardiac Arrest: a statement for healthcare professionals from a task force of the International Liaison Committee on Resuscitation (American Heart Association, European Resuscitation Council, Australian and New Zealand Council on Resuscitation, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Council of Southern Africa, Resuscitation Council of Asia); and the American Heart Association Emergency Cardiovascular Care Committee and the Council on Cardiopulmonary, Critical Care. Perioperative and Resuscitation Circulation. 2015;132:1286–300.
15. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. Crit Care Med. 1985;13:818–29.
16. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related organ failure Assessment) score to describe organ dysfunction/failure. On behalf of the working group on Sepsis-Related problems of the European society of intensive care medicine. Intensive Care Med. 1996;22:707–10.
17. Tibshirani R. Regression shrinkage and selection via the lasso: A retrospective. J Royal Stat Soc Ser B: Stat Methodol. 2011;73:273–82.
18. Huang W, Teo GKW, Tan JW, et al. Influence of comorbidities and clinical prediction model on neurological prognostication post out-of-hospital cardiac arrest. Heart Asia. 2018;10:e011016.
19. Nishioka N, Kobayashi D, Kiguchi T, et al. Development and validation of early prediction for neurological outcome at 90 days after return of spontaneous circulation in out-of-hospital cardiac arrest. Resuscitation. 2021;168:142–50.
20. Lin JJ, Huang CH, Chien YS, et al. TIMECARD score: an easily operated prediction model of unfavorable neurological outcomes in out-of-hospital cardiac arrest patients with targeted temperature management. J Formos Med Assoc. 2023;122:317–27.
21. Shimada T, Kawai R, Shintani A, et al. Neurological prognosis prediction upon arrival at the hospital after out-of-hospital cardiac arrest: R-EDByUS score. Resuscitation. 2024;200:110257.
22. Martinell L, Nielsen N, Herlitz J, et al. Early predictors of poor outcome after out-of-hospital cardiac arrest. Crit Care. 2017;21:96.
23. Oddo M, Sandroni C, Citerio G, et al. Quantitative versus standard pupillary light reflex for early prognostication in comatose cardiac arrest patients: an international prospective multicenter double-blinded study. Intensive Care Med. 2018;44:2102–11.
24. Choi JY, Jang JH, Lim YS, et al. Performance on the APACHE II, SAPS II, SOFA and the OHCA score of post-cardiac arrest patients treated with therapeutic hypothermia. PLoS ONE. 2018;13:e0196197.
25. Isenschmid C, Luescher T, Rasiah R, et al. Performance of clinical risk scores to predict mortality and neurological outcome in cardiac arrest patients. Resuscitation. 2019;136:21–9.
26. Lemoine L, Le Bastard Q, Batard E, et al. An Evidence-Based narrative review of the emergency department management of acute hyperkalemia. J Emerg Med. 2021;60:599–606.
27. Meng K, Lei X, He D. Association between serum calcium and in-hospital mortality in intensive care unit patients with cerebral infarction: a cohort study. Front Neurol. 2024;15:1428868.
28. Zheng X, Zhang F, Wang L, et al. Association between serum calcium and in-hospital mortality in critically ill atrial fibrillation patients from the MIMIC IV database. Sci Rep. 2024;14:27954.
29. Sun Y, He Z, Ren J, et al. Prediction model of in-hospital mortality in intensive care unit patients with cardiac arrest: a retrospective analysis of MIMIC -IV database based on machine learning. BMC Anesthesiol. 2023;23:178.
30. Hsu CH, Couper K, Nix T, et al. Calcium during cardiac arrest: A systematic review. Resusc Plus. 2023;14:100379.
31. Kikutani K, Nishikimi M, Shimatani T, et al. Differential effectiveness of hypothermic targeted temperature management according to the severity of Post-Cardiac arrest syndrome. J Clin Med. 2021;10:5643.
32. Lewis J, Perkins GD. Traumatic cardiac arrest. Curr Opin Crit Care. 2023;29:162–7.

Publisher's Note