Open Access Full Text Article

ORIGINAL RESEARCH

Tree-Based Algorithms and Association Rule Mining for Predicting Patients' Neurological Outcomes After First-Aid Treatment for an Out-of-Hospital Cardiac Arrest During COVID-19 Pandemic: Application of Data Mining

Wei-Chun Lin[1–3], Chien-Hsiung Huang[2–4], Liang-Tien Chien[4,5], Hsiao-Jung Tseng[6,7], Chip-Jin Ng [ID][2,3,8], Kuang-Hung Hsu [ID][2,3,9–12], Chi-Chun Lin[2,3,13], Cheng-Yu Chien [ID][2–4,13,14]

[1]Department of Emergency Medicine, New Taipei Municipal TuCheng Hospital and Chang Gung University, New Taipei City, Taiwan; [2]Department of Emergency Medicine, Chang Gung Memorial Hospital, Linkou and College of Medicine, Chang Gung University, Taoyuan, Taiwan; [3]Department of Emergency Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan; [4]Graduate Institute of Management, Chang Gung University, Taoyuan, Taiwan; [5]Fire Department, Taoyuan City Government, Taoyuan, Taiwan; [6]Department of Statistics and Information Science, Fu Jen Catholic University, New Taipei City, Taiwan; [7]Biostatistics Unit, Clinical Trial Center, Chang Gung Memorial Hospital, Taoyuan, Taiwan; [8]Department of Emergency Medicine, Chang Gung Memorial Hospital, Taipei, Taiwan; [9]Laboratory for Epidemiology, Chang Gung University, Taoyuan, Taiwan; [10]Laboratory for Epidemiology, Department of Health Care Management, Healthy Aging Research Center, Chang Gung University, Taoyuan, Taiwan; [11]Research Center for Food and Cosmetic Safety, College of Human Ecology, Chang Gung University of Science and Technology, Taoyuan, Taiwan; [12]Department of Safety, Health and Environmental Engineering, Ming Chi University of Technology, New Taipei City, Taiwan; [13]Department of Emergency Medicine, Ton-Yen General Hospital, Zhubei, Taiwan; [14]Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

Correspondence: Cheng-Yu Chien, Department of Emergency Medicine, Chang Gung Memorial Hospital, No. 5 Fushing St., Gueishan Dist, Taoyuan City, Taiwan, Tel +886-3-3281200 # 2505, Fax +886-3-3287715, Email rainccy217@gmail.com

Objective: The authors performed several tree-based algorithms and an association rules mining as data mining tools to find useful determinants for neurological outcomes in out-of-hospital cardiac arrest (OHCA) patients as well as to assess the effect of the first-aid and basic characteristics in the EMS system.

Patients and Methods: This was a retrospective cohort study. The outcome was Cerebral Performance Categories grading on OHCA patients at hospital discharge. Decision tree-based models inclusive of C4.5 algorithm, classification and regression tree and random forest were built to determine an OHCA patient's prognosis. Association rules mining was another data mining method which we used to find the combination of prognostic factors linked to the outcome.

Results: The total of 3520 patients were included in the final analysis. The mean age was 67.53 (±18.4) year-old and 63.4% were men. To overcome the imbalance outcome issue in machine learning, the random forest has a better predictive ability for OHCA patients in overall accuracy (91.19%), weighted precision (88.76%), weighted recall (91.20%) and F1 score (0.9) by oversampling adjustment. Under association rules mining, patients who had any witness on the spot when encountering OHCA or who had ever ROSC during first-aid would be highly correlated with good CPC prognosis.

Conclusion: The random forest has a better predictive ability for OHCA patients. This paper provides a role model applying several machine learning algorithms to the first-aid clinical assessment that will be promising combining with Artificial Intelligence for applying to emergency medical services.

Keywords: cardiac arrest, tree-based algorithms, data mining

Introduction

Out-of-hospital cardiac arrest (OHCA) poses a major threat to patient recovery because of the low rate of survival. This rate is approximately 2–10% in Asia and 9–15% in Europe and the United States. The survival at discharge rate of

patients with OHCA is nearly 7.6% in Taiwan.[1] Patients' survival after an OHCA depends on a series of actions enacted by bystanders, first responders, emergency medical service (EMS) personnel, and health-care staff (hospital staff). The chain of survival outlined by the American Heart Association includes early EMS activation, cardiopulmonary resuscitation, public automated external defibrillator use, high-quality resuscitation, and care after a cardiac arrest; this has been explored in several studies.[2–6] Factors such as the processing time for each stage, efficiency of emergency medical technicians (EMTs), and comorbidities of patients may help predict OHCA prognosis; however, few studies have investigated these factors. Thus, obtaining relevant data to improve survival rate and ensure favorable neurological outcomes in patients with OHCA remains a challenge.

Conventional statistical models, which can have some degree of sophistication, have been used to identify relevant factors and assign scores to predict OHCA prognosis. For instance, the cardiac arrest hospital prognosis score and C-GRApH scoring system are used for patients with return of spontaneous circulation (ROSC) after an OHCA to predict favorable neurological outcomes.[7,8] Because statistical models may be based on unrealistic assumptions for simplicity, a statistical model with high predictive ability may be overly complex. Owing to enhanced computational efficiency, machine learning (ML) techniques have been widely used in predictive modeling. In OHCA, various ML algorithms have been used to predict in-hospital cardiac arrest incidence and favorable neurological outcomes to identify an algorithm with enhanced performance and accuracy.[9–11] ML algorithms are used to classify patients with OHCA based on their care requirement after ROSC; target temperature management or percutaneous coronary intervention (PCI) reportedly lead to a better prognosis.[12] In Italy, a lower bystander CPR rate and prolonged response time reduced the survival of OHCA patients during the COVID-19 epidemic.[13] During the epidemic, the dispatch center identified OHCA, and EMTs conducted COVID-19 protection for all OHCA patients in Taiwan. Response time increased by 1 minute, scene time interval increased by 1.5 minutes, and t transport time increased by 0.5 minutes, which are also consistent with the previous study. However, bystander CPR decreased by 7%, and the rate of intubation and LMA decreased by 15%.[1] EMS parameters for OHCA patients have different situations to face the challenge of COVID-19.

Although ML-based models yield satisfactory predictions of survival or favorable neurological outcome in patients with OHCA, improper model application is sometimes unavoidable because EMS systems differ between countries.[14] Moreover, most studies have used ML techniques to emphasize prediction accuracy rather than interpretability. Therefore, in the present study, we used ML algorithms as data mining tools to identify the effective predictors of favorable neurological outcomes in patients with OHCA and evaluate the effects of first-aid treatment and basic EMS parameters on patient survival during the COVID-19 epidemic. We used several tree-based algorithms and association rule mining (ARM) to determine an optimal combination of the predictors of neurological outcomes in patients with OHCA.

Materials and Methods
Study Setting and Databases

This retrospective study was conducted using data obtained from the EMS database of Taoyuan, Taiwan, per the Utstein style between January 2020 and December 2021. Taoyuan, which has both urban and rural areas, has the third-highest rate of OHCA among municipalities in Taiwan. Our study obtained approval to access the EMS database utilized in the study. The resuscitation team must be dispatched with at least three EMTs and a mechanical cardiopulmonary resuscitation (CPR) device when attending to a patient with OHCA.

The EMS database is a local prospective registry maintained by the Taoyuan Fire department; the emergency departments of a total of 13 first-aid hospitals use this database to report cases. Patient data are collected by the EMS personnel and hospital staff (upon arrival at the emergency department). These data include the medical history of patients, age, sex, event location, witness status (ie, whether the incident was witnessed by others [positive or not]), CPR by bystanders, prehospital ROSC, initial rhythm (shockable or not), EMS parameters (EMT certification level, airway device, response time, scene time interval [STI], and transport time), outcome parameters evaluated using the Cerebral Performance Category (CPC) scale and survival at discharge. Furthermore, the hospitals' medical records included data on the patients' consciousness level and vital signs upon their arrival at the emergency department.

Patient Inclusion, Data Collection, and Outcome Measurement

The study included patients with nontraumatic OHCA who were managed and transported to the hospital by the EMS personnel. Patients were excluded if they had a do-not-resuscitate order, were pronounced dead at the scene, were pregnant, or had missing data on outcomes (Figure 1). The outcome assessed in this study was the neurological recovery of patients with OHCA. At discharge, the patients were evaluated using the CPC scale. Five grades can be assigned based on the CPC score. A favorable condition is indicated by a CPC grade of 1 (good cerebral performance and mild or no neurological disability) or 2 (moderate cerebral disability and consciousness and ability to function independently); a CPC grade of 5 indicates death. To simplify the evaluation of neurological recovery levels, three categories were defined based on CPC grading: good CPC (gCPC; CPC grades 1 and 2), severe cerebral damage (SCD; CPC grades 3 and 4), and death (CPC grade 5).

Data Processing

If a variable had missing values for <10% of its data points, these missing values were taken to be the mean for continuous and mode for categorical variables, respectively. If a variable had missing values for >10% of its data points, such as vital signs upon arrival at the hospital, it was excluded from the analysis. After the missing data were processed, the data were randomly split into training (80%) and testing (20%) data sets.

Statistical Analysis with Tree-Based ML

Decision tree–based models are common in supervised learning used for classification. For data analysis, we used three tree-based algorithms: C4.5, classification and regression tree (CART), and random forest (RF). The C4.5 algorithm follows the principle of maximum information entropy gain to help determine the best feature for splitting a dataset. It
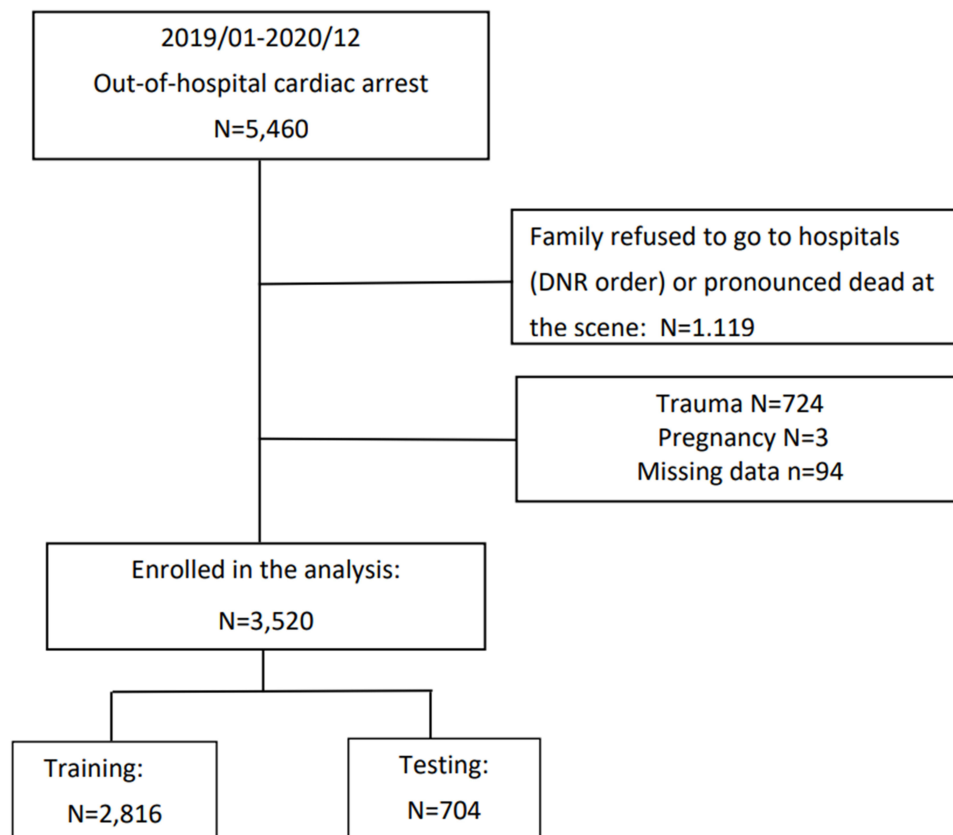

Figure 1 Flow diagram of patient enrollment.
Abbreviation: DNR, do-not-resuscitate.

can do multi-splits at once. CART is based on the Gini index, which indicates impurity. It only computes two splits at once. We respectively chose the optimal tree models with these two algorithms by achieving the highest accuracy in the testing dataset as well as lower pessimistic model error rate for reference. RF, as suggested by its name, comprises several decision trees and uses the bagging and random feature sampling techniques to reduce bias and increase precision for prediction. The other analysis is listed in Supplementary Figure S1 and S2.

Metrics for Model Performance Evaluation

To evaluate ML models, the accuracies of multicategory outcome (gCPC/ SCD/death) prediction were compared in terms of discrimination performance. The metrics used in this study included testing accuracy, kappa, multiclass area under the curve, precision, recall, and F1 score. The formula for these metrics are listed in Supplementary Table S1. Using RF, we ranked the prognostic factors based on their importance. The model performance in the three subgroups are listed in Supplementary Table S3.

Class Imbalance Problem

During our analysis, highly imbalanced outcome classes were noted in our data because of the low survival rate in patients with OHCA, which resulted in a high accuracy. To overcome this problem, data resampling was performed to retrain the models. Oversampling through the bootstrap method was applied to the tree-based ML algorithm described in the previous paragraph.

Other Statistical Analyses and ARM

Descriptive data for patient characteristics are presented as the mean ± standard deviation or median and interquartile range for continuous data; by contrast, frequency and percentage were used to present categorical data. Tree models help determine the possible paths for linking outcomes. Another data mining technique, ARM, was used to determine the rules for gCPC. An apriori algorithm was used to identify the rules for support, confidence, and lift of factor combinations. Downsampling was performed if no rules for gCPC could be identified using ARM. All statistical analyses were performed using R (version R i386 4.1.1). The main packages used for ML analyses were *RWeka, rpart, caret, e1071, randomForest*, and *pROC*.

Results

Among the patients with OHCA who were managed by the EMS personnel, 3520 were included in the final analysis. The mean age was 67.53 ± 18.4 years, and 63.4% were men. This cohort was divided into two samples: an 80% sample comprising a total of 2816 patients for training a predictive model for CPC-3 grade in patients with OHCA and the remaining 20% sample comprising a total of 704 patients for testing. The data on patient characteristics and their neurological outcomes are presented in Table 1. In the training cohort, the proportion of patients with gCPC was 4.7%, which was similar to that noted for all study participants, including the testing cohort. No significant differences were observed between the training and testing cohorts except in the following variables: CPR by bystanders and presence of hypertension and cancer in the patient.

Decision Tree

We used the C4.5 and CART algorithms to develop single-tree models. In C4.5 analysis, several combinations of the variables (ie, M = 2–15 with possible pruning during splitting; Supplementary Table S2) were used to determine the highest accuracy of the testing set. The highest accuracy that could be achieved using the C4.5 algorithm was 90.63% (parameter setting: M = 2, 3, and 13); however, the tree size was 79 with the lowest error rate (ie, M = 2), resulting in a more complex tree. In the CART analysis, we also considered the complexity penalty for pruning. Figure 2 shows the tree model developed using CART; the highest accuracy was 90.77% with a tree size of 13. The optimally pruned tree model decreased one split, where splitting stopped at the node of an STI of <16.5 min, remaining 90.34% for accuracy (Supplementary Figure S3). In RF analysis, top 10 factors were ranked based on their importance (Figure 3). The top five important factors were as follows: achievement of prehospital ROSC, age, EMS response time, STI, and transport time.

Table 1 Baseline Characteristics of the Study Population

	All OHCA N=3520	Training Set N=2816	Testing Set N=704	p-value
Age	67.53 ± 18.40	67.51 ± 18.35	67.62 ± 18.61	0.883
Sex=M	2232 (63.4%)	1801 (64.0%)	431 (61.2%)	0.178
Location				0.808
Public	787 (22.4%)	632 (22.4%)	155 (22.0%)	
Residential	2733 (77.6%)	2184 (77.6%)	549 (78.0%)	
Witness	1482 (42.1%)	1182 (42.0%)	300 (42.6%)	0.759
Bystander CPR	2271 (64.5%)	1841 (65.4%)	430 (61.1%)	0.033
Number of EMT members				0.190
1~2	446 (12.7%)	364 (12.9%)	82 (11.6%)	
3~4	2532 (71.9%)	2033 (72.2%)	499 (70.9%)	
5~6	542 (15.4%)	419 (14.9%)	123 (17.5%)	
Response time	6 (5 ~ 8)	6 (5 ~ 8)	6 (5 ~ 8)	0.921
Scene time interval	13 (10 ~ 17)	13 (10 ~ 17)	13 (10 ~ 17)	0.183
Transport time	6 (4 ~ 9)	6 (4 ~ 9)	6 (4 ~ 9)	0.155
EMS total time	27 (23 ~ 32)	27 (23 ~ 32)	26 (22 ~ 31)	0.099
Use mechanical CPR	2930 (83.2%)	2340 (83.1%)	590 (83.8%)	0.652
Airway				0.723
BVM	2656 (75.5%)	2119 (75.2%)	537 (76.3%)	
Igel/LMA	626 (17.8%)	502 (17.8%)	124 (17.6%)	
Intubation	238 (6.8%)	195 (6.9%)	43 (6.1%)	
AED Rhythm				0.848
Non-shockable	3008 (85.5%)	2408 (85.5%)	600 (85.2%)	
Shockable	512 (14.5%)	408 (14.5%)	104 (14.8%)	
Epinephrine medication times				0.181
0	2736 (77.7%)	2182 (77.5%)	554 (78.7%)	
1	226 (6.4%)	186 (6.6%)	40 (5.7%)	
2	231 (6.6%)	181 (6.4%)	50 (7.1%)	
3	205 (5.8%)	160 (5.7%)	45 (6.4%)	
4	122 (3.5%)	107 (3.8%)	15 (2.1%)	
Amiodarone medication (1+)	26 (0.7%)	21 (0.7%)	5 (0.7%)	0.922
Glasgow Coma Scale				0.834
3	3452 (98.1%)	2760 (98.0%)	692 (98.3%)	
4~8	20 (0.6%)	17 (0.6%)	3 (0.4%)	
9~15	48 (1.4%)	39 (1.4%)	9 (1.3%)	
Patients' History				
DM	924 (26.3%)	741 (26.3%)	183 (26.0%)	0.863
HTN	1260 (35.8%)	1038 (36.9%)	222 (31.5%)	0.008
Stroke	266 (7.6%)	215 (7.6%)	51 (7.2%)	0.726
CAD	865 (24.6%)	709 (25.2%)	156 (22.2%)	0.096
Cancer	337 (9.6%)	285 (10.1%)	52 (7.4%)	0.027
CPC category				0.280
Good CPC (1 2)	160 (4.5%)	131 (4.7%)	29 (4.1%)	
Severe Cerebral Damage (3 4)	205 (5.8%)	172 (6.1%)	33 (4.7%)	
Death (5)	3155 (89.6%)	2513 (89.2%)	642 (91.2%)	

Abbreviations: CPR, cardiopulmonary resuscitation; EMS, emergency medical service; BVM, bag valve mask; LMA, laryngeal mask airway; AED, automated external defibrillator; DM, diabetes mellitus; HTN, hypertension; CAD, coronary artery disease; CPC, cerebral performance category.

Model Performance

Table 2 shows the overall performance of models developed using the three tree-based algorithms. No statistically significant difference was noted among the algorithms in terms of the model performance metrics. In the oversampling
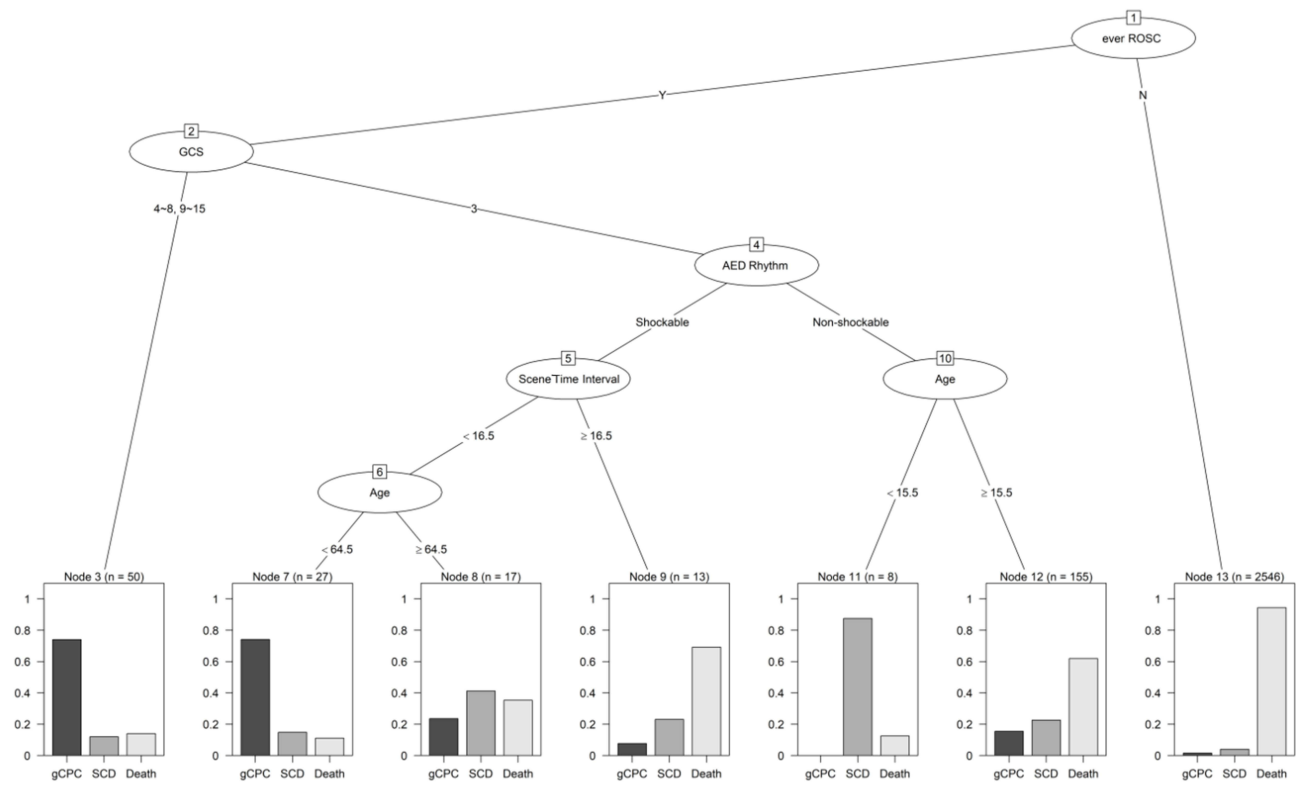
Figure 2 Tree model plot by classification and regression tree (CART).
Abbreviations: gCPC, Good Cerebral Performance Category; SCD, severe cerebral damage; GCS, Glasgow coma scale; ROSC, return of spontaneous circulation; AED, automated external defibrillator.
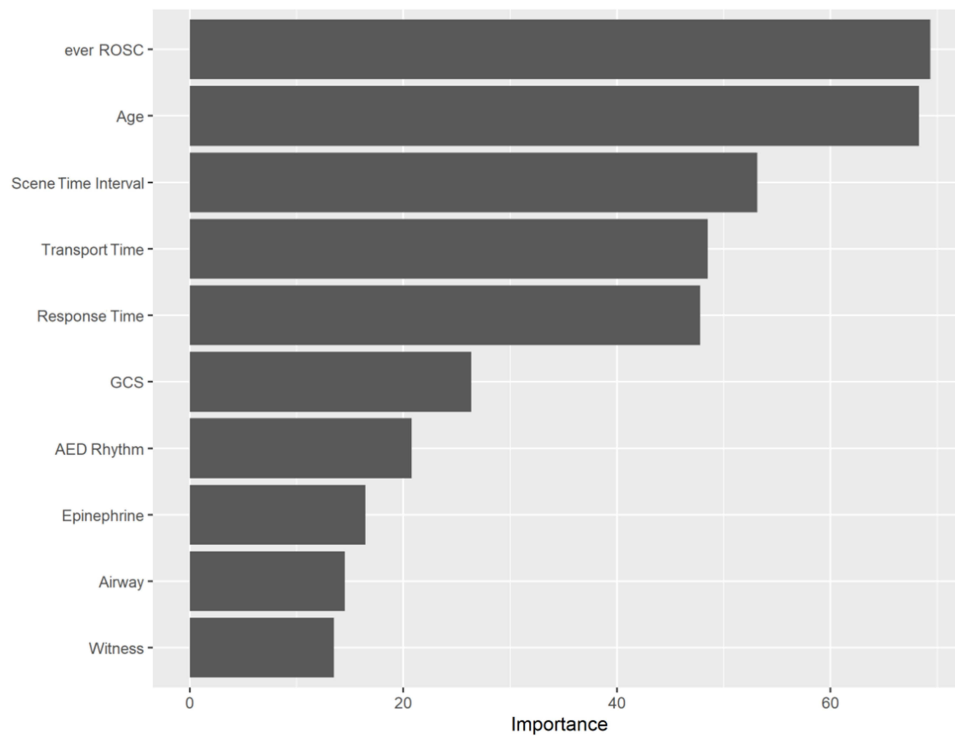

Figure 3 The top 10 of parameter importance ranked by the random forest (RF).
Abbreviations: ROSC, return of spontaneous circulation; GCS, Glasgow coma scale; AED, automated external defibrillator.

Table 2 Evaluation Results for Tree-Based Model Performances

Overall Performance	C4.5		CART		RF	
	Original Dataset	Over-Sampling	Original Dataset	Over-Sampling	Original Dataset	Over-Sampling
Accuracy	90.62%	81.68%	90.77%	80.82%	90.91%	91.19%
Kappa	0.384	0.221	0.353	0.328	0.341	0.434
Multi-class AUC	64.67%	70.19%	64.46%	76.66%	65.15%	71.47%
Weighted Precision	88.55%	85.30%	90.77%	80.82%	88.13%	88.76%
Weighted Recall	90.63%	81.68%	88.67%	88.01%	90.91%	91.20%
F1 score	0.896	0.834	0.897	0.843	0.895	0.900

Note: Weighted metrics are calculated by the summation of the proportion (subgroup numbers/ total) multiplying the subgroups' metrics.
Abbreviations: CART, classification and regression tree; RF, random forest.

Table 3 Rules for Good CPC by 1:1 Resampling for Association Rules Mining

{Conditions} ⊆{Outcome}	Support	Confidence	Lift
{Witness=Y} ≤ {CPC=gCPC}	0.425	0.720	1.439
{Ever ROSC=Y} ≤ {CPC=gCPC}	0.344	0.917	1.833
{Mechanical CPR =N} ≤ {CPC=gCPC}	0.203	0.730	1.461
{Witness=Y, Cancer=N, Epinephrine=0} ≤ {CPC=gCPC}	0.338	0.715	1.430
{Gender=M, Witness=Y} ≤ {CPC=gCPC}	0.309	0.712	1.424
{AED Rhythm=Shockable} ≤ {CPC=gCPC}	0.275	0.807	1.615
{Location=Public} ≤ {CPC=gCPC}	0.250	0.755	1.509
{Hospital Level=CAC, Witness=Y} ≤ {CPC=gCPC}	0.244	0.736	1.472
{Response Time=1~5 mins, Witness=Y} ≤ {CPC=gCPC}	0.219	0.700	1.400
{Witness=Y, Bystander CPR=Y, Airway=BVM, Cancer=N} ≤ {CPC=gCPC}	0.219	0.707	1.414
{Age=45~65 y/o, Witness=Y} ≤ {CPC=gCPC}	0.209	0.761	1.523

Notes: Model setting: min support=0.2, min confidence=0.7; listing in the raw output with original variable names from the dataset.
Abbreviations: gCPC, Good Cerebral Performance Category; ROSC, return of spontaneous circulation; CAC, cardiac arrest center; AED, automated external defibrillator; BVM, bag valve mask; CPR, cardiopulmonary resuscitation.

analysis, the metrics of the model developed using RF were slightly higher than those of the models developed using the other algorithms. Specifically, the overall accuracy, weighted precision, weighted recall, and F1 score were 91.19%, 88.76%, 91.20%, and 0.9, respectively. Thus, RF had a better predictive ability for CPC-3 grade in patients with OHCA.

ARM

The apriori algorithm was used to determine the rules for gCPC in patients with OHCA. However, all rules led to death because of the highly imbalanced distribution of outcomes (Supplementary Table S4). Therefore, we performed 1:1 resampling to determine the outcomes of patients with gCPC versus the others. The rules determined under the conditions of a minimum support value of 0.2 and a minimum confidence value of 0.7 are presented in Table 3. Event witness status or prehospital ROSC were highly correlated with good OHCA prognosis. For instance, a patient with OHCA for whom the EMS response time is 1–5 min and whose OHCA event is witnessed by others may have a favorable prognosis.

Discussion

We used several data mining techniques to develop an optimal model for predicting patients' neurological outcomes after an OHCA. RF was used to rank the prognostic factors based on their importance. In our study, the model developed using RF had better accuracy, precision, and recall than the others, particularly in the oversampling analysis, for survival and

favorable neurological outcomes after an OHCA. The top five important prognostic factors identified were prehospital ROSC, age, EMS response time, STI, and transport time.

Patients with ROSC after an OHCA were reported to have a better (almost eight-fold) chance of survival at discharge than those without it.[15] In our study, this factor represented the first important branch node in the decision tree. In contrast to the OHCA patients without ROSC, it has >80% mortality rate.

A previous study reported that the average duration from call to hospital arrival was approximately 30 min for gCPC; every 1-min increase in the response time was associated with a 3% decrease in the probability of favorable neurological outcomes.[16] Through ARM, we found that the presence of a witness at the scene and a response time of <5 min lead to favorable neurological outcomes in patients with OHCA. Similarly, another study reported that a response time of <4 min results in a better ROSC rate and survival in the first year after the event. Moreover, prolonged CPR (>30 min) results in poor prognosis.[17] Holmén et al showed that patients for whom the response time is <6 min have a survival probability 6% and 11% higher than that of those for whom the response time is 7–9 and >15 min, respectively. If the patient's rhythm is nonshockable, a prolonged response time leads to a decrease in survival probability by 10%–19%.[18]

We found that patients with OHCA with a shockable rhythm achieved prehospital ROSC. However, they were still comatose and had an on-spot first-aid time of >16.5 min (STI > 16.5); these patients had a lower (one-fifth fold) probability of achieving favorable neurological outcomes than those with an on-spot first-aid time of <16.5 min. Coute et al reported that the probability of achieving favorable neurological outcomes rapidly decreases in patients with OHCA with a shockable rhythm when the STI is >20 min; however, to the best of our knowledge, no such finding has been reported in patients with a nonshockable rhythm. This finding is consistent with our finding obtained using the decision tree–based model.[19]

In their retrospective study conducted in South Korea using the data of patients with OHCA, Kim et al concluded that a prolonged STI leads to poor prognosis in patients with OHCA; the cutoff point reported in their study was 8 min, which is shorter than the node split observed in our model developed using the CART algorithm.[19] They further stated that if an EMT performs on-spot CPR for >8 min in a patient without ROSC, the probability of survival is low after arrival at the hospital.[20]

Chien demonstrated that the prognosis of patients with OHCA with a shockable rhythm who are directly transported to a cardiac arrest center is not affected the transport time. However, when the transport time is >14 min, the probability of survival is lower than the average.[1] Patients receive definitive treatments, such as target temperature management, PCI, and extracorporeal membrane oxygenation (ECMO)-CPR, at a cardiac arrest center.[1] In their meta-analysis, Geri et al reported that prolonged transport time is not correlated with survival and neurological outcomes.[21] Nonetheless, transport time is an important factor as per the ranking based on RF. Currently, we lack clarity on the suitable transport time (to definitive care hospitals) for patients with OHCA.

In our study, among comatose patients with ROSC, with a shockable rhythm, and with an STI of <16.5 min, patients aged <65 years exhibited better neurological outcomes in CART analysis. ARM suggested that a combination of factors such as middle age and positive witness status leads to favorable neurological outcomes. In a study by Okabayashi et al, the patient population was divided into three groups based on age: 65–74 [reference group], 75–84, and >85 years; the probability of favorable neurological outcomes was 46% and 69% lower in patients aged 75–84 and 85 years, respectively, than in those aged 65–74 years.[22] However, on the basis of an age-based analysis, Tsai et al[2] reported only a 2–3% difference in terms of survival and neurological outcomes among the age groups, which was not as significant as that reported by Okabayashi.[22] In another previous study, age was included as a confounding factor in outcome evaluation, but the findings remained nonsignificant.[1]

Multiple factors are correlated with favorable neurological outcomes, including witness status and the patient's initial shockable rhythm. In our resampling analysis, a positive witness status and response time of >5 min was correlated with a favorable neurological outcome of 0.700. This finding is consistent with our results related to the important prognostic factors ranked using RF. A positive witness status and bag-valve-mask ventilation performed by bystanders to rescue patients with OHCA without cancer was associated with a favorable neurological outcome of 0.707 in our study. Reynolds et al also demonstrated that in patients with OHCA whose cardiac event is witnessed by others, who receive CPR from the bystanders, and who have a modified Rankin Scale score of 0–3 had a two-fold higher survival than those

whose cardiac event is unwitnessed and who receive no CPR from the bystanders.[23,24] In the Çinier's study, primary percutaneous coronary intervention (PPCI) for STEMI patients had a longer ischemic time and door-to-balloon time prolonged by 28 minutes, which does not affect in-hospital mortality risk.[25] In the COVID-19 pandemic, the EMS total time was extended by 3 minutes, but the survival to discharge rate was similar.

Because the probability of survival is low after an OHCA, highly imbalanced outcome categories may result in high accuracy because of the incorporation of highly correlated factors in the model; however, these factors do not reflect the model's advantages. Resampling techniques can be used to evaluate these ML techniques. Single-tree models can be developed using the C4.5 and CART algorithms, which helps determine the rules that can predict patients' neurological outcomes. Although CART is used more often than C4.5 algorithm in medical studies, both are similar in terms of overall accuracy (81.68% vs 80.82%, respectively; resampling analysis). Conversely, RF is an ensemble learning approach, which combines multiple weak learners to construct a robust model. The top 10 prognostic factors ranked in our study based on their importance may serve as a reference for relevant studies in the future. Due to artificial intelligence (AI) technology, we should use AI to predict disease prognosis. Hayıroğlu's study used the electrocardiographic diastolic index (EDI) to create a new scoring system to indicate diastolic dysfunction (DD).[26] Our study identified critical prognostic factors using Decision tree-based models and association rules mining. We should use survival factors to build scoring systems to predict the survival rate in OHCA patients during the covid-19 pandemic.

Strengths and Limitations

We evaluated factors associated with neurological recovery in patients with OHCA to evaluate the effects of first-aid treatment and basic EMS parameters on patient survival. However, no in-hospital factors, such as ECMO-CPR, PCI, and target temperature management, were analyzed in the present study. Hence, further studies are warranted in the future to combine in-hospital and out-of-hospital data for predicative models developed based on tree-based algorithms and ARM.

Conclusions

RF has a better predictive ability for CPC-3 grade in patients with OHCA. In decision tree, ARM, and RF analyses, the key prognostic factors in an out-of-hospital setting were prehospital ROSC, age, response time, STI, and transport time. The model developed in this study using several ML algorithms to evaluate the effects of first-aid treatment may be combined with artificial intelligence to enhance the EMS system.

Abbreviations

OHCA, out-of-hospital cardiac arrest; EMS, emergency medical service; EMTs, emergency medical technicians; CPR, cardiopulmonary resuscitation; ML, machine learning; ECMO, extracorporeal membrane oxygenation; PCI, percutaneous coronary intervention; BVM, bag valve mask; LMA, laryngeal mask airway; AED, automated external defibrillator; DM, diabetes mellitus; HTN, hypertension; CAD, coronary artery disease; CPC, Cerebral Performance Category; gCPC, Good Cerebral Performance Category; ROSC, return of spontaneous circulation; CAC, cardiac arrest center; BVM, bag valve mask; AI, artificial intelligence; EDI, electrocardiographic diastolic index; dd, diastolic dysfunction; GCS, Glasgow coma scale.

Data Sharing Statement

All data relevant to the study are included in the article or uploaded as <u>Supplementary Information</u>.

Ethics Approval and Consent to Participate

The data accessed complied with relevant data protection and privacy regulations. This study was reviewed and approved by the Institutional Review Board of Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan (approval no. 202200559B0). The requirement for informed consent was waived owing to the retrospective nature of this study. Our study obtained approval to access the EMS database utilized in the study.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Chien CY, Tsai SL, Tsai LH, et al. Impact of transport time and cardiac arrest centers on the neurological outcome after out-of-hospital cardiac arrest: a retrospective cohort study. *J Am Heart Assoc*. 2020;9:e015544. doi:10.1161/JAHA.119.015544

2. Tsai SL, Chaou CH, Huang CH, et al. Features of hospital and emergency medical service in out-of-hospital cardiac arrest patients with shockable rhythm. *Am J Emerg Med*. 2017;35:1222–1227. doi:10.1016/j.ajem.2017.03.032

3. Huang CH, Fan HJ, Chien CY, et al. Validation of a dispatch protocol with continuous quality control for cardiac arrest: a before-and-after study at a city fire department-based dispatch center. *J Emerg Med*. 2017;53:697–707. doi:10.1016/j.jemermed.2017.06.028

4. Chien CY, Chien WC, Tsai LH, et al. Impact of the caller's emotional state and cooperation on out-of-hospital cardiac arrest recognition and dispatcher-assisted cardiopulmonary resuscitation. *Emerg Med J*. 2019;36:595–600. doi:10.1136/emermed-2018-208353

5. Chien CY, Fang SY, Tsai LH, et al. Traditional versus blended CPR training program: a randomized controlled non-inferiority study. *Sci Rep*. 2020;10:10032. doi:10.1038/s41598-020-67193-1

6. Chien CY, Weng YM, Hsu SC, Kuo CW, Chaou CH. Effect of population-based training programs on bystander willingness to perform cardiopulmonary resuscitation. *J Intensive Care Emerg Med*. 2016;12:63–69.

7. Sauneuf B, Dupeyrat J, Souloy X, et al. The CAHP (cardiac arrest hospital prognosis) score: a tool for risk stratification after out-of-hospital cardiac arrest in elderly patients. *Resuscitation*. 2020;148:200–206. doi:10.1016/j.resuscitation.2020.01.011

8. Kiehl EL, Parker AM, Matar RM, et al. C-GRApH: a validated scoring system for early stratification of neurologic outcome after out-of-hospital cardiac arrest treated with targeted temperature management. *J Am Heart Assoc*. 2017;6:e003821. doi:10.1161/JAHA.116.003821

9. Liu N, Koh ZX, Goh J, et al. Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection. *BMC Medical Inform Decis Mak*. 2014;14:75. doi:10.1186/1472-6947-14-75

10. Kwon JM, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. *J Am Heart Assoc*. 2018;7:e008678. doi:10.1161/JAHA.118.008678

11. Park JH, Shin SD, Song KJ, et al. Prediction of good neurological recovery after out-of-hospital cardiac arrest: a machine learning analysis. *Resuscitation*. 2019;142:127–135. doi:10.1016/j.resuscitation.2019.07.020

12. Harford S, Darabi H, Del Rios M, et al. A machine learning based model for out of hospital cardiac arrest outcome classification and sensitivity analysis. *Resuscitation*. 2019;138:134–140. doi:10.1016/j.resuscitation.2019.03.012

13. Uy-Evanado A, Chugh HS, Sargsyan A, et al. Out-of-hospital cardiac arrest response and outcomes during the COVID-19 pandemic. *JACC Clin Electrophysiol*. 2021;7(1):6–11. doi:10.1016/j.jacep.2020.08.010

14. Ge Y, Lou Y, Xu M, et al. Spatial distribution and influencing factors on the variation of bacterial communities in an urban river sediment. *Environ Pollut*. 2021;272:115984. doi:10.1016/j.envpol.2020.115984

15. Lee DE, Lee MJ, Ahn JY, et al. New termination-of-resuscitation models and prognostication in out-of-hospital cardiac arrest using electrocardiogram rhythms documented in the field and the emergency department. *J Korean Med Sci*. 2019;34:e134. doi:10.3346/jkms.2019.34.e134

16. Onoe A, Kajino K, Daya MR, et al. Outcomes of patients with OHCA of presumed cardiac etiology that did not achieve prehospital restoration of spontaneous circulation: the all-Japan Utstein Registry experience. *Resuscitation*. 2021;162:245–250. doi:10.1016/j.resuscitation.2021.03.003

17. Sladjana A, Gordana P, Ana S. Emergency response time after out-of-hospital cardiac arrest. *Eur J Intern Med*. 2011;22:386–393. doi:10.1016/j.ejim.2011.04.003

18. Holmén J, Herlitz J, Ricksten SE, et al. Shortening ambulance response time increases survival in out-of-hospital cardiac arrest. *J Am Heart Assoc*. 2020;9:e017048. doi:10.1161/JAHA.120.017048

19. Coute RA, Nathanson BH, Kurz MC, McNally B, Mader TJ; CARES Surveillance Group. The association between scene time interval and neurologic outcome following adult bystander witnessed out-of-hospital cardiac arrest. *Am J Emerg Med*. 2021;46:628–633. doi:10.1016/j.ajem.2020.11.059

20. Kim KH, Shin SD, Song KJ, et al. Scene time interval and good neurological recovery in out-of-hospital cardiac arrest. *Am J Emerg Med*. 2017;35:1682–1690. doi:10.1016/j.ajem.2017.05.049

21. Geri G, Gilgan J, Wu W, et al. Does transport time of out-of-hospital cardiac arrest patients matter? A systematic review and meta-analysis. *Resuscitation*. 2017;115:96–101. doi:10.1016/j.resuscitation.2017.04.003

22. Okabayashi S, Matsuyama T, Kitamura T, et al. Outcomes of patients 65 years or older after out-of-hospital cardiac arrest based on location of cardiac arrest in Japan. *JAMA Netw Open*. 2019;2:e191011. doi:10.1001/jamanetworkopen.2019.1011

23. Reynolds JC, Grunau BE, Rittenberger JC, Sawyer KN, Kurz MC, Callaway CW. Association between duration of resuscitation and favorable outcome after out-of-hospital cardiac arrest: implications for prolonging or terminating resuscitation. *Circulation*. 2016;134:2084–2094. doi:10.1161/CIRCULATIONAHA.116.023309

24. Morgan DP, Muscatello D, Hayen A, Travaglia J. Human factors influencing out-of-hospital cardiac arrest survival. *Emerg Med Australas*. 2019;31:600–604. doi:10.1111/1742-6723.13222

25. Çinier G, Hayıroğlu M, Pay L, et al. Effect of the COVID-19 pandemic on access to primary percutaneous coronary intervention for ST-segment elevation myocardial infarction. *Turk Kardiyol Dern Ars*. 2020;48(7):640–645. doi:10.5543/tkda.2020.95845

26. Hayıroğlu Mİ, Çınar T, Çiçek V, et al. A simple formula to predict echocardiographic diastolic dysfunction—electrocardiographic diastolic index. *Herz*. 2021;46(Suppl 2):159–165. doi:10.1007/s00059-020-04972-6