Available online at www.sciencedirect.com

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Clinical paper

# Deep-learning-based out-of-hospital cardiac arrest prognostic system to predict clinical outcomes

*Joon-myoung Kwon*[a,1], *Ki-Hyun Jeon*[b,*,1], *Hyue Mee Kim*[b], *Min Jeong Kim*[b], *Sungmin Lim*[b], *Kyung-Hee Kim*[b], *Pil Sang Song*[b], *Jinsik Park*[b], *Rak Kyeong Choi*[b], *Byung-Hee Oh*[b]

[a] *Department of Emergency Medicine, Mediplex Sejong Hospital, Incheon, Republic of Korea*
[b] *Division of Cardiology, Cardiovascular Center, Mediplex Sejong Hospital, Incheon, Republic of Korea*

## Abstract

**Aim:** Out-of-hospital cardiac arrest (OHCA) is a major healthcare burden, and prognosis is critical in decision-making for treatment and the withdrawal of life-sustaining therapy. This study aimed to develop and validate a deep-learning-based out-of-hospital cardiac arrest prognostic system (DCAPS) for predicting neurologic recovery and survival to discharge.

**Methods:** The study subjects were patients from the Korea OHCA registry who experienced return of spontaneous circulation (ROSC) after OHCA. A total of 36,190 patients were exclusively divided into a set of 28,045 subjects for derivation data and 8,145 subjects for validation data. We used information available for the time of ROSC as predictor variables, and the endpoints were neurologic recovery (cerebral performance category 1 or 2) and survival to discharge. The DCAPS was developed using the derivation data and represented the favorability of prognosis with a score between 0 and 100.

**Results:** The area under the receiver operating characteristic curve (AUROC) of DCAPS for predicting neurologic recovery for the validation data was 0.953 [95% confidence interval 0.952–0.954]; these results significantly outperformed those of logistic regression (0.947 [0.943–0.948]), random forest (0.943 [0.942–0.945]), support vector machine (0.930 [0.929–0.932]), and conventional methods of a previous study (0.817 [0.815–0.820]). The AUROC of the DCAPS for survival to discharge was 0.901 [0.900–0.903], and this result significantly outperformed those of other models as well.

**Conclusions:** The DCAPS predicted neurologic recovery and survival to discharge of OHCA patients accurately and outperformed the conventional method and other machine-learning methods.

**Keywords:** Artificial intelligence, Out-of-Hospital cardiac arrest, Prognosis, Neural networks, Machine learning, Decision support techniques

*Abbreviations:* AUROC, area under the receiver operating characteristic curve; CPC, cerebral performance category; CPR, cardiopulmonary resuscitation; DCAPS, deep-learning-based out-of-hospital cardiac arrest prognostic system; ECG, electrocardiography; ED, emergency departments; EMS, emergency medical service; KOHCAR, Korea Out-of-hospital cardiac arrest registry; LR, logistic regression; OHCA, out-of-hospital cardiac arrest; RF, random forest; ROSC, return of spontaneous circulation; SVM, support vector machine.

* Corresponding author at: Division of Cardiology, Department of Internal Medicine, Cardiovascular Center, Mediplex Sejong Hospital, 20, Gyeyangmunhwa-ro, Gyeyang-gu, Incheon, Republic of Korea.
E-mail address: imcardio@gmail.com (K.-H. Jeon).

[1] These two authors contributed equally to this study.

https://doi.org/10.1016/j.resuscitation.2019.04.007
Received 26 January 2019; Received in revised form 16 March 2019; Accepted 3 April 2019

## Introduction

Out-of-hospital cardiac arrest (OHCA) is a major healthcare burden worldwide.[1,2] Among those with return of spontaneous circulation (ROSC), the in-hospital mortality rate is 50–70%, and the majority sustain neurological injury.[1–3] Prediction of prognosis is critical in decision-making for the treatment of patients with OHCA and is an important factor in the withdrawal of life sustaining therapy (WLST) during the post-resuscitation period. Furthermore, prediction of prognosis-based decision making for using mechanical circulatory support (e.g., extra-corporeal membrane oxygenation, tandem heart, or impella) or for performing early coronary angiography that could improve the clinical outcomes of patients were important in the initial resuscitation.[4]

Several prognostic models have been developed in previous studies.[3,5,6] However, the performances of these models were not accurate enough to determine treatment or WLST. These models used inaccurate variables dependent on memory during a highly stressful time, and variables that could be obtained after admission. And these models used only limited information based on a conventional statistical approach, such as multivariate analysis by the logistic regression, which is potentially limited by information loss.[7,8]

In order to develop an accurate prognostic system for OHCA using only initial information, we used deep-learning. Recently, deep-learning has shown high performance in several medical domains, such as image detection and clinical outcome prediction.[9–11] In this study, we developed and validated a deep-learning-based out-of-hospital cardiac arrest prognostic system (DCAPS). To the best of our knowledge, this study is the first to predict prognosis of OHCA using deep-learning.

## Methods

### Study population

We conducted a retrospective observation cohort study using data from the Korea OHCA Registry (KOHCAR) in South Korea. OHCA surveillance was conducted by Korean centers for disease control and prevention (CDC), and all 712 emergency departments (ED) were included. The full details of the KOHCAR aims and protocols have been published elsewhere.[12] The Korea CDC approved the use of all data. The institutional review boards of Sejong General Hospital approved this study and granted waivers of informed consent based on general impracticability and minimal harm. Patient information was anonymized and de-identified before being provided.

We included patients who experienced ROSC after initial resuscitation and excluded patients with missing values in this study. We defined ROSC as persistent circulation for at least 20 min after return of spontaneous circulation.

### Predictor variables and endpoints

The study data was split exclusively into derivation and validation datasets according to the year. The derivation data was the patient data for 2012–2015, and the validation data was the patient data for 2016.

The DCAPS is a system that predicts the clinical outcomes of each patient after ROSC. We used only the information available at the time of ROSC, including *age, sex, place of OHCA, etiology of OHCA* (disease or trauma), *ROSC in emergency medical service (EMS)* (ROSC before ED visit), *event witness, bystander cardiopulmonary resuscitation (CPR), initial electrocardiography (ECG) rhythm of EMS, initial ECG rhythm of ED*, and *time from ED visit to* ROSC, as predictor variables (Table 1). If the patient had a positive value for ROSC in EMS, we regarded the value *of time from ED visit to ROSC* of the patient as -1.

We did not use information that depended on memory during a highly stressful situation and may have been inaccurate (e.g. last normal time, witnessed time, no flow time, and low flow time). We did not use information such as laboratory test results that were not available at the time of initial resuscitation. The primary endpoint was neurologic recovery (CPC 1 or 2) after ROSC, and the secondary endpoint was survival to discharge after ROSC.

### Development of a deep-learning-based prognostic system

As shown in Fig. 1, we developed the DCAPS using only the derivation data. The DCAPS was developed using a multilayer perceptron (MLP), a method of deep-learning.[13] Because there was no gain in accuracy when adding more than 5 layers, we used 5 layers to minimize the parameters to be learned. We used a rectified linear unit (ReLU) as the activation function[14] and used TensorFlow (the Google Brain Team) as the backend.[15] Furthermore, we used the Adagrad optimizer with default parameters, and binary cross-entropy as the loss function.[16] Before using the derivation data for the DCAPS development, we conducted data preprocessing.[17] To train the model, we used the backpropagation method.[13,17] We provided a more detailed description of deep-learning methods in Supplemental material 1.

### Development of machine-learning based prognostic models

We also developed three machine-learning models—logistic regression (LR), support vector machine (SVM), and random forest (RF)—for performance comparison with the DCAPS.[18,19] In previous studies, LR, SVM, and RF have been the most typically used machine-learning methods and have shown better performance than traditional methods in several medical domains.[20,21] We provided the detail description and architecture of each machine-learning methods as Supplemental material 1.

### Validation of prognostic system

After we developed the DCAPS, we confirmed the performance of the models using only the validation data that were not used for model development. We used the area under the receiver operating characteristic curve (AUROC) as the measure. The AUROC is a frequently used metric, and the receiver operating characteristic (ROC) curve plots sensitivity against (1–specificity).[22] We evaluated the 95% confidence interval using bootstrapping (10,000 sampling iterations with replacement).[23] We also compared the performance of DCAPS with the derived machine-learning and conventional prognostic models described by Goto et al., who developed a decision-tree model to predict outcome after out-of-hospital cardiac arrest in the emergency department (Supplemental material 2).[24] That model used only the definite variables which were initially identifiable using only the history and were less dependent on uncertain memory. Because the characteristics of the variables were similar to those used in the

**Table 1 – Baseline characteristics of study subjects[a].**

| Variable | All (n = 36,190) | Derivation data (n = 28,045) | Validation data (n = 8145) | *P*-value[*] |
|---|---|---|---|---|
| Age (year, mean ± SD) | 61.88 ± 18.13 | 61.61 ± 18.18 | 62.80 ± 17.95 | <0.001 |
| Female (n, %) | 12,236 (33.81) | 9497 (33.86%) | 2739 (33.63%) | 0.702 |
| Witnessed | 22,398 (61.89%) | 16,978 (60.54%) | 5420 (66.54%) | <0.001 |
| ROSC in EMS | 4526 (12.51%) | 3159 (11.26%) | 1367 (16.78%) | <0.001 |
| Bystander CPR | 7690 (21.25%) | 4692 (16.73%) | 2998 (36.81%) | <0.001 |
| Place of OHCA | | | | <0.001 |
| Home | 16,440 (45.43%) | 12,803 (45.65%) | 3637 (44.65%) | |
| Public space | 7867 (21.74%) | 6176 (22.02%) | 1691 (20.76%) | |
| Ambulance | 3443 (9.51%) | 2521 (8.99%) | 922 (11.32%) | |
| Others | 8440 (23.32%) | 6545 (23.34%) | 1895 (23.27%) | |
| Trauma | 8010 (22.13%) | 6306 (22.49%) | 1704 (20.92%) | 0.003 |
| Initial ECG rhythm of EMS | | | | <0.001 |
| VF | 4388 (12.12%) | 3174 (11.32%) | 1214 (14.90%) | |
| Pulseless VT | 391 (1.08%) | 314 (1.12%) | 77 (0.95%) | |
| Asystole | 8062 (22.28%) | 5728 (20.42%) | 2334 (28.66%) | |
| PEA | 23349 (64.52%) | 18,829 (67.14%) | 4520 (55.49%) | |
| Initial ECG rhythm of ED | | | | <0.001 |
| VF | 1316 (3.6%) | 1149 (4.10%) | 167 (2.05%) | |
| Pulseless VT | 281 (0.78%) | 242 (0.86%) | 39 (0.48%) | |
| Asystole | 17151 (47.39%) | 13,344 (47.58%) | 3807 (46.74%) | |
| PEA | 17442 (48.20%) | 13,310 (47.46%) | 4132 (50.73%) | |
| ED visit to ROSC time (min, mean ± SD) | 20.47 ± 22.66 | 20.98 ± 23.05 | 18.73 ± 21.17 | <0.001 |

[a] CPR denotes cardiopulmonary resuscitation, ECG electrocardiography, ED emergency department, EMS emergency medical services, PEA pulseless electrical activity, ROSC return of spontaneous circulation, SD standard deviation, VF ventricular fibrillation, and VT ventricular tachycardia.
[*] The alternative hypothesis for this *p*-value was that there is a difference between the derivation and validation data groups for each variable.

DCAPS, we used Goto's predictive model as the conventional method for comparison.

We compared the difference between prognostic models. We confirmed the importance of variables in the derivation data for each model. And we grouped patients as high and low risk with each prognostic models. The cutoff point of each model was confirmed when the sensitivity was approximately 95%. This is the highest sensitivity value seen in conventional models. The prognostic model for OHCA patient needs high sensitivity because there would be no error to cause termination of treatment for patients with a good prognosis.

## Results

We included 36,795 OHCA patients from the KOHCAR who experienced ROSC between January 2012 and December 2016 and excluded 605 patients owing to missing data. The study subjects were comprised of 36,190 patients, where 3,812 had neurologic recovery, and 7,157 experienced survival to discharge. The DCAPS was developed using 28,045 patients from the derivation data. An accuracy test was performed using the validation data of 8,145 patients, where 977 had neurologic recovery, and 1,719 experienced survival to discharge. The patients' characteristics in the derivation and validation datasets were described in Table 1. We provide the developed DCAPS, coding book for making input tidy data, example of tidy data, and python code for validation as a supplement files to this article.

The AUROC of DCAPS, LR, RF, SVM, and conventional model was 0.976. 0.955, 0.965, 0.951, and 0.821 in the training dataset for predicting neurologic recovery; further, the AUROC of DCAPS, LR, RF, SVM, and conventional model was 0.928, 0.891, 0.886, 0.856, and 0.734 in training dataset for predicting survival to discharge, respectively. As shown in Fig. 2, during the validation, the AUROC of the DCAPS was 0.953(95% confidence interval(CI), 0.952–0.954] for neurologic recovery (primary endpoint). This result significantly outperformed the LR (0.947[0.946–0.948]), RF (0.943[0.942–0.945]), and SVM (0.930[0.929–0.932]). The AUROC of the conventional model was 0.817[0.815–0.820], which is the same as the accuracy given in the previous study. For the secondary endpoint, survival to discharge, the AUROC of the DCAPS was 0.901[95% CI 0.900–0.903]. This result significantly outperformed the other models.

After we predicted the endpoint with each prognostic model using the cutoff, we confirmed the sensitivity, specificity, positive predictive value, negative predictive value, F-measure, and net reclassification index (NRI) as shown in Fig. 3. As we fixed the sensitivities when comparing accuracy, the number of patients with a correctly predicted neurologic recovery was similar for each model, as shown in Venn-diagram of Fig. 3. The DCAPS model correctly predicted the non-neurologic recovery patients more than the other models and its accuracy was the most consistent in all metrics. Compared with the DCAPS, the NRI of the machine learning models was significantly positive, which means that the DCAPS was more accurate than the other methods.[25]

As shown in Supplemental material 3, the variable importance is different for each prognosis model. For other machine learning models, the *ROSC in EMS* and *ED visit to ROSC time* variables are important to prediction. However, for the DCAPS, the *ROSC in EMS* variable is less important to prediction. To confirm the reason for this, we conducted an additional analysis, as shown in Supplemental material 4. The material shows how the importance of the variables changes when the deep-learning-based model is re-developed after removing the most important variables. The material shows the correlation matrix between predictor variables. After we removed the *ED visit to ROSC time* variable, the *ROSC in*

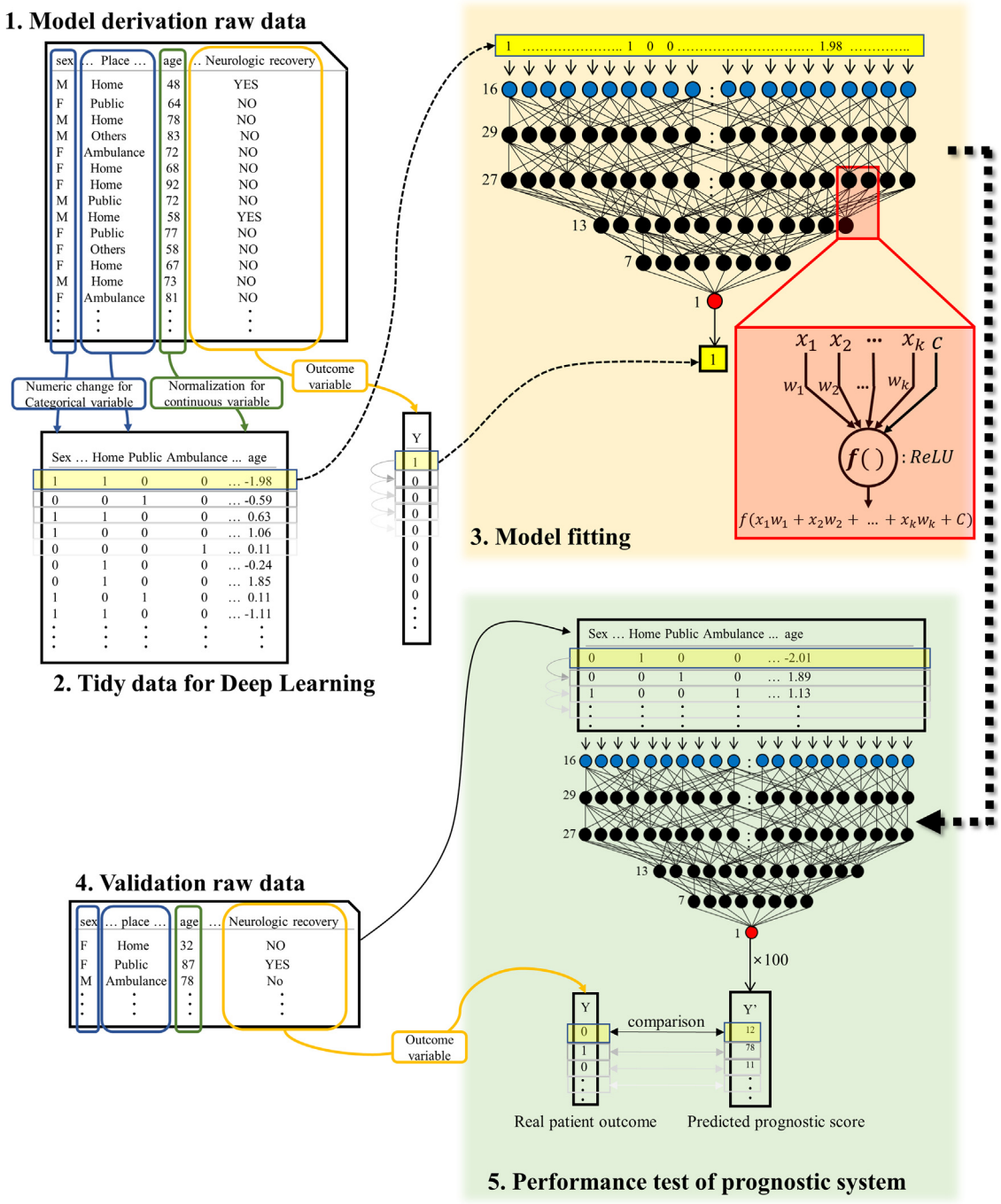


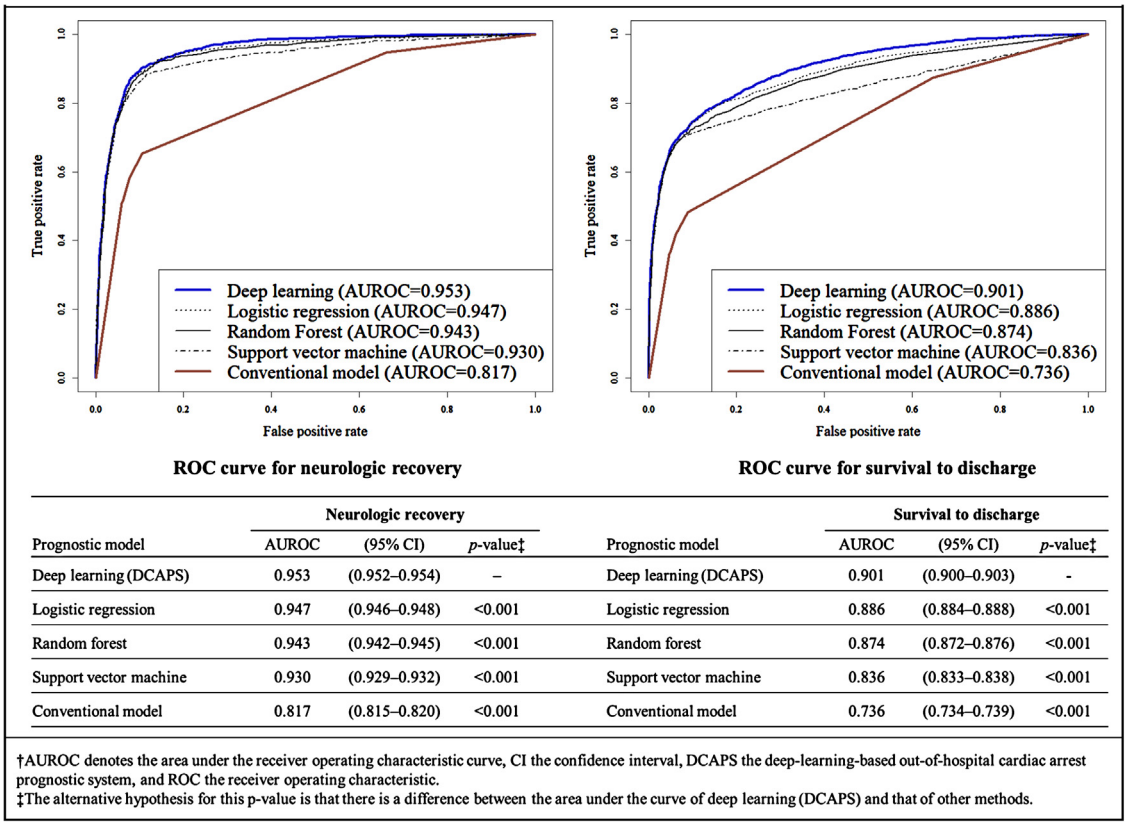

**Fig. 1 – †ReLU denotes the rectified linear unit. A more detailed description of deep-learning methods was provided in Supplemental material 1.**

*EMS* became the most important variable in the deep-learning model.

## Discussion

In this study, we developed a prognostic prediction model for the clinical outcomes of patients after OHCA using deep-learning on data from a large national registry. Through validation, this study demonstrated that the accurate performance of the deep-learning model, DCAPS, was excellent for predicting neurologic recovery and survival to discharge. The most important finding from this study was that the models derived from artificial intelligence were better than conventional model to predict clinical outcomes of OHCA patients and deep-learning method was the best among them.

When we apply DCAPS to electronic health records (EHR) in a hospital and EMS, the possibility of neurologic recovery could be calculated in real time. In this setting, we could make decisions such as




| | Neurologic recovery | | | | Survival to discharge | | |
|---|---|---|---|---|---|---|---|
| Prognostic model | AUROC | (95% CI) | *p*-value‡ | Prognostic model | AUROC | (95% CI) | *p*-value‡ |
| Deep learning (DCAPS) | 0.953 | (0.952–0.954) | – | Deep learning (DCAPS) | 0.901 | (0.900–0.903) | - |
| Logistic regression | 0.947 | (0.946–0.948) | <0.001 | Logistic regression | 0.886 | (0.884–0.888) | <0.001 |
| Random forest | 0.943 | (0.942–0.945) | <0.001 | Random forest | 0.874 | (0.872–0.876) | <0.001 |
| Support vector machine | 0.930 | (0.929–0.932) | <0.001 | Support vector machine | 0.836 | (0.833–0.838) | <0.001 |
| Conventional model | 0.817 | (0.815–0.820) | <0.001 | Conventional model | 0.736 | (0.734–0.739) | <0.001 |

†AUROC denotes the area under the receiver operating characteristic curve, CI the confidence interval, DCAPS the deep-learning-based out-of-hospital cardiac arrest prognostic system, and ROC the receiver operating characteristic.
‡The alternative hypothesis for this p-value is that there is a difference between the area under the curve of deep learning (DCAPS) and that of other methods.

**Fig. 2 – †AUROC denotes the area under the receiver operating characteristic curve, CI the confidence interval, DCAPS the deep-learning-based out-of-hospital cardiac arrest prognostic system, and ROC the receiver operating characteristic. ‡The alternative hypothesis for this p-value is that there is a difference between the area under the curve of deep learning (DCAPS) and that of other methods.**

sustaining the effort of resuscitation. Further, we could determine invasive treatments, such as coronary angiography, and life-sustaining efforts using mechanical circulatory support, such as extra-corporeal membrane oxygenation, tandem heart, and impella. Further, we could be guided to determine the withdrawal of life-sustaining therapy (WLST).

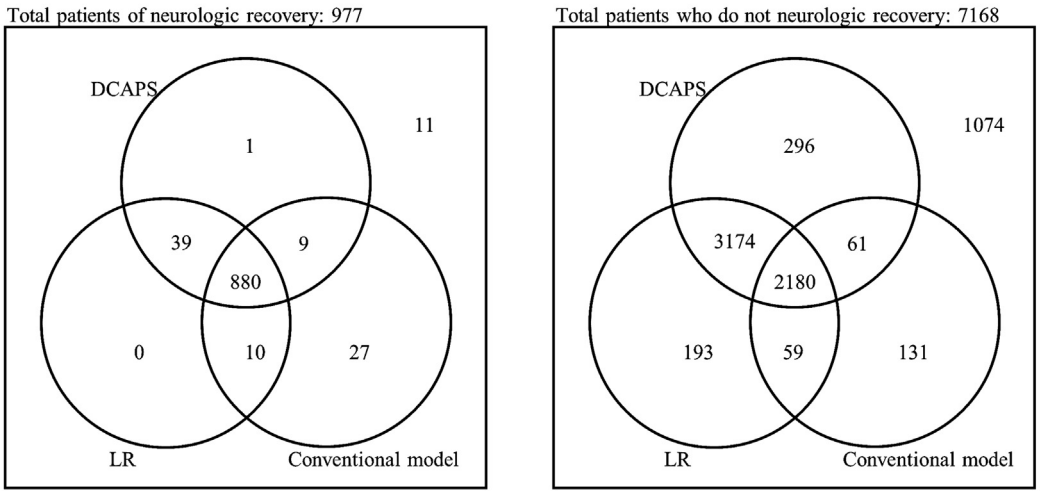
The factors that predict the mortality of OHCA have been reported previously, such as age, immediate bystander resuscitation, ROSC, clearly treatable cardiac cause, no-flow time, and initial shockable rhythm.[4,26] Temperature management treatment is also known as a predictor of good neurologic outcomes, despite the controversy.[27] Because the etiologies of OHCA were varied, and many unpredictable factors were included in this clinical setting, prediction of prognosis was complex and not one-size-fits-all. Maupain et al. suggested the Cardiac Arrest Hospital Prognosis (CAHP) score from a large-multicentric population-based cohort, which identified seven independent predictors including age, initial non-shockable rhythm, home setting arrest, delays between collapse and BLS and between BLS and ROSC, dose of epinephrine, and acidosis for poor neurologic outcome.[5] Although CAHP is well developed and validated, the score calculation is quite complex for cardiac arrest and ROSC situations, and requires variables that are not available at initial resuscitation, such as laboratory test results. The DCAPS has good accuracy for predicting prognosis and requires only information that is available at the time of ROSC.

In this study, DCAPS showed a very high accuracy that outperformed the results of previous studies.[3,5,6] The models of previous studies using LR has limitations, including the fixed assumptions on data behavior, and the need to preselect variables in the development phase, thus leading to potential information loss.[7,8] Additionally, old models cannot represent the relationship between variables because the score is measured only by the sum of the variables. Meanwhile, deep-learning can use all available information without potential loss[8,28] and obtains the relationship between the variables, as shown in Fig. 1, unlike conventional methods.

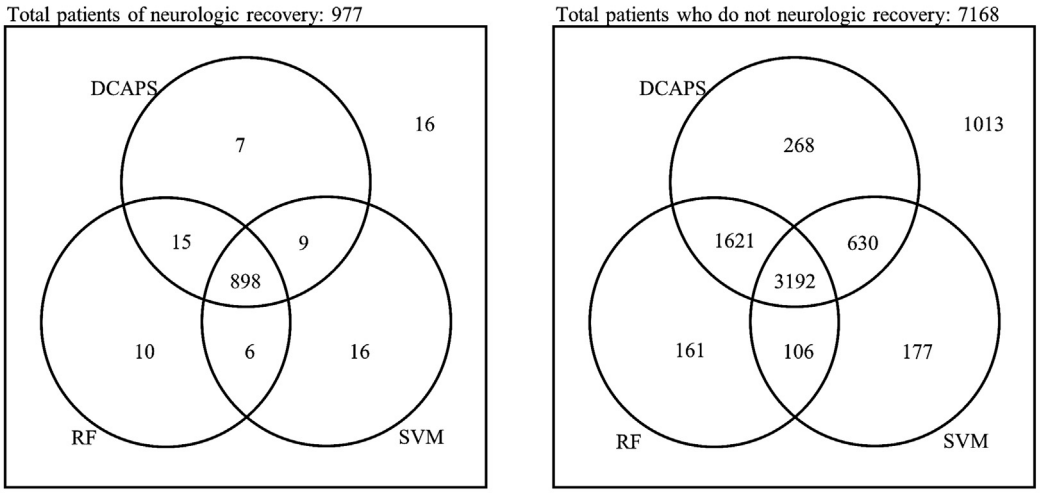
The previous studies attempted to predict survival to discharge after ROSC using conventional machine-learning methods, including LR, SVM, and RF.[21,29] These conventional machine-learning techniques were limited in their ability to process natural data in their raw form. The machine-learning required careful engineering to design a feature extractor that transforms the raw data into suitable internal representation. In other words, the researcher must not only choose a variable but also decide which relationship of the many between the variables to use. This process requires a lot of manpower, and there are possibilities of important information being missed.[12]

Deep-learning includes feature-learning which is a set of methods that allow a model to be fed with raw data and to automatically identify the features needed for conducting a task. Deep-learning comprises multiple processing layers of feature, obtained by composing simple but non-linear modules, each of which transforms the feature at one level (starting with the raw input) into a feature at a higher, slightly more abstract level. Higher layers of presentation not only amplify the

**A. Venn diagram for comparison between DCAPS, logistic regression model, and conventional model**




**B. Venn diagram for comparison between DCAPS, support vector machine model, and random forest model**




| Prognostic model | **Sensitivity** (95% CI) | **Specificity** (95% CI) | **PPV** (95% CI) | **NPV** (95% CI) | **F-measure** (95% CI) | **NRI ‡** (95% CI) |
|---|---|---|---|---|---|---|
| **Deep learning (DCAPS)** | 0.951 (0.948-0.954) | 0.797 (0.795-0.799) | 0.389 (0.385-0.394) | 0.992 (0.991-0.992) | 0.552 (0.548-0.557) | - |
| **Logistic regression** | 0.951 (0.948-0.954) | 0.782 (0.780-0.784) | 0.373 (0.368-0.377) | 0.992 (0.991-0.992) | 0.535 (0.531-0.540) | 0.0146 (0.0057-0.0258) |
| **Random forest** | 0.951 (0.948-0.954) | 0.709 (0.706-0.711) | 0.308 (0.304-0.312) | 0.991 (0.990-0.991) | 0.465 (0.461-0.470) | 0.0880 (0.0733-0.1029) |
| **Support vector machine** | 0.951 (0.948-0.954) | 0.573 (0.570-0.575) | 0.233 (0.229-0.235) | 0.988 (0.988-0.989) | 0.374 (0.370-0.377) | 0.2241 (0.2063-0.2398) |
| **Conventional model** | 0.948 (0.945-0.951) | 0.339 (0.337-0.342) | 0.164 (0.161-0.166) | 0.979 (0.978-0.981) | 0.279 (0.276-0.282) | 0.1993 (0.1691-0.2321) |

† CI denotes the confidence interval, DCAPS the deep-learning-based out-of-hospital cardiac arrest prognostic system, LR logistic regression, NPV negative predictive value, RF random forest, SVM support vector machine, NRI net reclassification index, and PPV positive predictive value.
‡ Net reclassification index = (net correctly reclassified neurologic recovery patients using DCAPS / number of neurologic recovery patients) + (net correctly reclassified patients who did not have neurologic recovery using DCAPS / number of patients who did not have neurologic recovery)

**Fig. 3 – †CI denotes the confidence interval, DCAPS the deep-learning-based out-of-hospital cardiac arrest prognostic system, LR logistic regression, NPV negative predictive value, NRI net reclassification index, and PPV positive predictive value. ‡ Net reclassification index = (net correctly reclassified neurologic recovery patients using DCAPS/ number of neurologic recovery patients) + (net correctly reclassified patients who did not have neurologic recovery using DCAPS/number of patients who did not have neurologic recovery).**

aspects of the input that are important for conducting a task but also suppress irrelevant variations.[12] It is important to note that these layers of features of deep-learning are not designed by humans. As this process is conducted automatically, it is good at discovering the intricate structures in high-dimensional data without information loss and it requires only very little engineering by humans. Therefore, it can be quickly applied to other tasks with ease. Owing to these prospects, deep-learning is applied in various domains of medicine, and it shows better performance than all the other methods.[9,28]

Compared with the importance of the *ROSC in EMS* variable in the LR, RF, and SVM models, the *ROSC in EMS* variable was less important in DCAPS. As shown in Supplemental material 4, after we removed the *ED visit to ROSC time* variable, the *ROSC in EMS* became the most important variable in the deep-learning model. This result suggested that the deep-learning method has the ability to represent the relationship between variables. As the odds ratio derived from LR has been used when assessing risk factors for many diseases, the fact that deep-learning can represent the relationship between variables can help in the investigation of new risk factors for disease.

As shown in Table 1, there were several significant differences between training and validation datasets. A standardized dispatcher-assisted bystander cardiopulmonary resuscitation was implemented nationwide in 2015. And many efforts of basic life support education for lay person were the reason that there was different between derivation and validation datasets. Machine-learning and deep-learning models memorize the characteristics of the derivation dataset. Many previous machine-learning and deep-learning studies in the medical field divided the study data with randomization to generate a separated validation dataset. Using this method, however, there was little difference in character between derivation and validation data in many studies. And the accuracy of predictive model could be overestimated than the accuracy of real world. To solve this problem, we split the data according to time and defined the validation dataset as that closet to the present. Since the purpose of the algorithms in to predict the future, we believed this to be a reasonable approach. And it was an important point that the DCAPS predicted endpoints accurately in the validation dataset which had different characteristics from those of derivation data.

Several limitations are present in our study. First, deep-learning is known as a *black box*. Although we can fit the deep-learning model by confirming each weight, we cannot interpret the deep-learning model in terms of the approach to the decision for clinical outcome. Explainable deep-learning has been studied recently, and this will be our next area of study.[29] Second, it is necessary to validate this model in the global OHCA population because the dataset used in the present study included only South Korean patients. For this reason, we have provided as supplemental files our prediction model, data preprocessing methods, and code for validation. On the other hand, as mentioned above, it is easy to create a deep learning algorithm, so researchers can develop algorithms for each region based on this research. Finally, because data for the study population was nationwide data from the Korea CDC, detailed characteristics for each patient, such as body mass index, comorbidity, and medications, were not available.

## Conclusion

In conclusion, we developed and validated a new prognostic model for OHCA based on the deep-learning approach. The DCAPS predicted the neurologic recovery and survival to discharge of OHCA patients more accurately than other machine-learning and conventional methods. This study showed the potential and effectiveness of a deep-learning-based algorithm model for resuscitation, which can be applied to electronic health records and will be a useful and fast tool to identify patients with a good neurologic outcome and help precise decision-making in daily practice.

## Funding source

No funding was secured for this study Financial Disclosure: The all authors have no financial relationships relevant to this article to disclose

## Disclosures

All authors have no conflicts of interest.

## Acknowledgement

None.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.resuscitation.2019.04.007.

REFERENCES

1. Gräsner J-T, Lefering R, Koster RW, et al. EuReCa ONE-27 Nations, ONE Europe, ONE Registry. Resuscitation 2016;105:188–95.
2. Benjamin EJ, Virani SS, Callaway CW, et al. Heart disease and stroke statistics—2018 update: a report from the American heart association. Circulation 2018.
3. Adrie C, Cariou A, Mourvillier B, et al. Predicting survival with good neurological recovery at hospital admission after successful resuscitation of out-of-hospital cardiac arrest: The OHCA score. Eur Heart J 2006;27:2840–5.
4. Hassager C, Nagao K, Hildick-Smith D. Out-of-hospital cardiac arrest: in-hospital intervention strategies. Lancet Elsevier Ltd 2018;391:989–98.
5. 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.
6. Martinell L, Nielsen N, Herlitz J, et al. Early predictors of poor outcome after out-of-hospital cardiac arrest. Crit Care Critical Care 2017;21:1–10.
7. Sun G, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. J Clin Epidemiol 1996;49:907–16.
8. Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). Stat Sci 2001;16:199–231.
9. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 2018;15:20170387.
10. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama 2016;304:649–56.

11. Kwon J-M, Lee Y, Lee Y, et al. An algorithm based on deep learning for predicting in-hospital cardiac arrest. J Am Heart Assoc 2018;7: e008678.
12. Kim YT, Do Shin S, Hong SO, et al. Effect of national implementation of utstein recommendation from the global resuscitation alliance on ten steps to improve outcomes from Out-of-Hospital cardiac arrest: a ten-year observational study in Korea. BMJ Open 2017;7:e016925.
13. Pal SK, Mitra S. Multilayer perceptron, fuzzy sets, and classification. IEEE Trans Neural Networks 1992;3:683–97.
14. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. Proceeding of the 27th International Conference Machine Learning. . p. 807–14.
15. Abadi M, Barham P, Chen J, et al. TensorFlow: a system for large-scale machine learning tensorflow: a system for large-scale machine learning. 12th USENIX Symposium Operating System Design Implement (OSDI' 16). . p. 265–84.
16. Kingma DP, Ba J. Adam: a method for stochastic optimization. 2017 IEEE International Conference Consumer Electronics ICCE 2017. . p. 434–5.
17. Jayalakshmi T, Santhakumaran A. Statistical normalization and backpropagation for classification. Int J Comput Theory Eng 2011;3:89–93.
18. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak 2011;11:51 BioMed Central Ltd.
19. Calcagno V, De Mazancourt C. Glmulti: an r package for easy automated model selection with (Generalized) linear models. J Stat Softw 2010;34:1–29.
20. Churpek MM, Yuen TC, Winslow C, et al. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit Care Med 2016;44:368–74.
21. Krizmaric M, Verlic M, Stiglic G, et al. Intelligent analysis in predicting outcome of out-of-hospital cardiac arrest. Comput Methods Programs Biomed 200995: Elsevier Ireland Ltd.
22. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27:861–74.
23. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 2000;19:1141–64.
24. Goto Y, Maeda T, Goto Y. Decision-tree model for predicting outcomes after out-of-hospital cardiac arrest in the emergency department. Crit Care 2013;17:R133.
25. Leening MJG, Vedder MM, Witteman JCM, et al. Net reclassification improvement: computation, interpretation, and controversies. Ann Intern Med 2014;160:122–31.
26. Hawkes C, Booth S, Ji C, et al. Epidemiology and outcomes from out-of-hospital cardiac arrests in England. Resuscitation 2017;110:133–40.
27. Bernard S, Gray T, Buist M, et al. Treatment of comatose survivors of out-of-hospital cardiac arrest with induced hypothermia. N Engl J Med 2002;346:557–63.
28. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
29. Fong RC, Vedaldi A. Interpretable explanations of black boxes by meaningful perturbation. Proceeding IEEE International Conference on Computer Vision 2017 2017;3449–57.