Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

RESUSCITATION 178 (2022) 55–62

Available online at ScienceDirect

Resuscitation

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

Predicting neurological outcomes after in-hospital cardiac arrests for patients with Coronavirus Disease 2019

Anoop Mayampurath[a,b], Fereshteh Bashiri[b], Raffi Hagopian[c], Laura Venable[d], Kyle Carey[d], Dana Edelson[d], Matthew Churpek[a,b,], for the American Heart Association's Get With The Guidelines®-Resuscitation Investigators*

Abstract

Background: Machine learning models are more accurate than standard tools for predicting neurological outcomes in patients resuscitated after cardiac arrest. However, their accuracy in patients with Coronavirus Disease 2019 (COVID-19) is unknown. Therefore, we compared their performance in a cohort of cardiac arrest patients with COVID-19.

Methods: We conducted a retrospective analysis of resuscitation survivors in the Get With The Guidelines®-Resuscitation (GWTG-R) COVID-19 registry between February 2020 and May 2021. The primary outcome was a favorable neurological outcome, indicated by a discharge Cerebral Performance Category score ≤ 2. Pre- and peri-arrest variables were used as predictors. We applied our published logistic regression, neural network, and gradient boosted machine models developed in patients without COVID-19 to the COVID-19 cohort. We also updated the neural network model using transfer learning. Performance was compared between models and the Cardiac Arrest Survival Post-Resuscitation In-Hospital (CASPRI) score.

Results: Among the 4,125 patients with COVID-19 included in the analysis, 484 (12 %) patients survived with favorable neurological outcomes. The gradient boosted machine, trained on non-COVID-19 patients was the best performing model for predicting neurological outcomes in COVID-19 patients, significantly better than the CASPRI score (c-statistic: 0.75 vs 0.67, $P < 0.001$). While calibration improved for the neural network with transfer learning, it did not surpass the gradient boosted machine in terms of discrimination.

Conclusion: Our gradient boosted machine model developed in non-COVID patients had high discrimination and adequate calibration in COVID-19 resuscitation survivors and may provide clinicians with important information for these patients.

Keywords: Cardiac arrest, Prediction, Neurological outcomes, Machine learning

Introduction

Accurate prognostication of neurological status in survivors of in-hospital cardiac arrest (IHCA) is essential for patient families, as it informs decision-making regarding goals of care and could be valuable for risk standardization and quality improvement initiatives.[1–3] However, prognostication of neurological status is challenging for resuscitation survivors because these patients are often intubated, sedated, and in a state of induced hypothermia. Therefore, researchers have developed tools, such as the Cardiac Arrest Survival Post-Resuscitation In-Hospital (CASPRI) score, to predict the likelihood of favorable neurological outcomes at discharge using pre- and peri-arrest variables.[4]

In prior work using a cohort derived from the Get With the Guidelines Resuscitation (GWTG-R) registry from 2009 to 2017, we demonstrated that an extreme gradient boosted (XGBoost) machine learning model predicted favorable neurological status at discharge significantly better than the CASPRI score.[5] The XGBoost model also outperformed all other machine learning models, such as the logistic regression (LR) and the multi-layer perceptron (MLP) neural network, in terms of discrimination, calibration, and accuracy measures. However, the number of Coronavirus Disease 2019 (COVID-19) cases remains high across the United States, and neurological prognostication in resuscitated patients with COVID-19 involves additional challenges. Recent studies have reported low survival rates among COVID-19 patients who experience cardiac

* Corresponding author at: 8007 Excelsior Drive, Madison, WI 53717-1903, United States.
E-mail address: mchurpek@medicine.wisc.edu (M. Churpek).

https://doi.org/10.1016/j.resuscitation.2022.07.018
Received 22 April 2022; Received in Revised form 13 July 2022; Accepted 14 July 2022

arrests.[6–10] Resuscitation survivors also have a poor likelihood of being discharged without neurological deficits.[6,7,11] Additionally, the increased risk of exposure for care personnel, the requirement of personal protective equipment, and the shortage of staff and supplies impact resuscitation practice and assessment of neurological status.[12,13] With these factors, it is unknown how previously published models that predict neurological outcome perform among COVID-19 resuscitation survivors.

Therefore, the aim of this study was to validate the performance of CASPRI and our prior machine learning models for predicting favorable neurological outcomes in a cohort of resuscitation survivors with COVID-19. We hypothesized that our machine learning models derived from resuscitated patients without COVID-19 would predict favorable neurological status at discharge in resuscitated COVID-19 patients more accurately than CASPRI. We further hypothesized that we could utilize transfer learning, a machine learning framework that updates a previously developed model in a new dataset, to improve the performance of our MLP neural network model in predicting favorable neurological outcomes in COVID-19 patients.

Methods

Data sources and study population

We accessed the GWTG-R COVID-19 registry to build our COVID-19 study population. Hospitals participating in the registry submit clinical information regarding the medical history, care, and outcomes of consecutive patients hospitalized for in-hospital cardiac arrest using an online, interactive case report form and Patient Management Tool™ (IQVIA, Parsippany, New Jersey). We identified 11,173 in-hospital cardiac arrests within the GWTG-R COVID-19 registry (see **Supplementary** Fig. 1) corresponding to patients with confirmed or suspected COVID-19 between February 2020 and May 2021. We utilized the same exclusion criteria as our recent study,[5] eliminating subsequent cardiac arrests for an individual patient (n = 2,078), removing arrests outside of general medicine or intensive care unit (ICU) settings (n = 1,122), patients without recorded return of spontaneous circulation (n = 3,600), missing discharge survival status (n = 134), or missing Cerebral Performance Category (CPC) assessment on discharge (n = 114). The institutional review board at the University of Wisconsin-Madison reviewed and approved the study with a waiver of informed consent (IRB# 2020-0588).

Primary outcome and predictors

The primary outcome of interest was a favorable neurological outcome at the time of patient discharge, defined as a CPC score of ≤ 2, per the outcome definition of the CASPRI score and our previously developed machine learning models.[4,5] We retained the same set of predictors as our models and CASPRI, which include pre- and peri-arrest variables related to patient and arrest characteristics, neurologic status prior to arrest, pre-existing conditions, and interventions in place prior to arrest.

Model development

Our prior study demonstrated that the LR, XGBoost, and MLP models were top-performing in terms of discrimination, calibration, and accuracy metrics.[5] We thus retrained these models using the entire cohort of 117,383 patients without COVID-19 (combined training and testing data from our published study). Missing values were addressed depending on the algorithm. Data for the LR and MLP models were imputed using predictions from decision trees created from the non-COVID-19 derivation dataset. Briefly, we created classification (for categorical features) or regression (for numeric variables) decision trees from complete non-missing observations within the non-COVID-19 derivation dataset. These trees were then used to predict impute values for missing data in the COVID-19 dataset. The XGBoost was trained with missing values, as the algorithm can natively handle missing data.

We further created a new machine learning model using transfer learning to specialize our retrained MLP model to adapt to patients in our COVID-19 cohort. Briefly, transfer learning is a machine learning technique wherein the weights of all layers of a neural network except for the last are frozen, and then additional layers with trainable weights are added. Thus, the foundational layers are trained to recognize broad patterns for predicting the initial outcome, while the final layers are trained to predict a different outcome or the same outcome in a different patient population.[14] In our study, we added a trainable dense layer to our original MLP architecture. Thus, the initial layers of this new model, called MLP-Transfer, were already trained to detect global features to predict neurological outcomes in a general non-COVID-19 population of resuscitation survivors, while the new final dense layer is explicitly trained to predict neurological outcomes in COVID-19 survivors of resuscitation. We employed a nested cross-validation approach to train MLP-Transfer (see **Supplementary** Fig. 2). Briefly, we divided the COVID-19 population into five folds. Data from four folds were used to train the MLP-Transfer model with an 80%–20% derivation-validation split for hyperparameter optimization, while the fifth fold formed the independent test set. This strategy was iterated five times, after which prediction probabilities for all observations were concatenated to assess model performance. This strategy enabled direct comparison between our models in this study for the same number of test observations. We utilized the caret package in R Version 3.6.0 (R Project for Statistical Computing) for training the LR and XGB models and the keras and keras_tuner packages in Python 2.7 for training and optimizing hyperparameters for the MLP and MLP-Transfer models.

Model performance

Our primary metric to assess model performance was the discrimination of COVID-19 resuscitation survivors with favorable neurological outcomes at discharge, as indicated by the area under the receiver operating characteristic curve (AUC). We compared model AUCs to each other and to CASPRI using DeLong's method.[15] Model calibration was assessed by calculating unreliability index (U), indicating divergence between log-likelihood of the uncalibrated and calibrated response variable, and by testing for H0: intercept = 0, slope = 1.[16] We further calculated sensitivity, specificity, negative and positive predictive values for the best-performing machine learning models and the CASPRI score. A cutoff of $P < 0.05$ was used to indicate statistical significance. Finally, we estimated variable importance using a permutation-based method that equates feature importance for predicting favorable neurological outcome with loss function changes when the feature is permuted. We also report performance in accordance with the Transparent Reporting of multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines (checklist in **Supplementary** Table 1).

Results

Patient characteristics

Among the 4,125 patients that met our inclusion criteria from 241 hospitals, 484 (12 %) patients survived with favorable neurological outcomes, indicated by a CPC score ≤ 2 at discharge. Comparisons of patient and arrest characteristics between COVID-19 resuscitated patients with and without survival with a favorable neurological outcome are shown in Table 1. COVID-19 patients with favorable neurological outcomes were younger (mean age: 60 vs 65 years, $P < 0.001$), had a shorter duration of arrest (median time: 5 vs 8 minutes, $P < 0.001$), had higher use of AED (51 % vs 43 %, $P = 0.002$), and lower CPC score prior to arrest (CPC Score 1: 72 % vs 58 %, CPC Score 2: 18 % vs 15 %, $P < 0.001$) compared to those without favorable neurological outcome. The median length of stay for patients in our cohort was 13 days (IQR: 6–23 days). Table 2 compares the rate of pre-existing conditions and pre-arrest interventions between COVID-19 resuscitation survivors with and without our primary outcome. Patients with favorable neurological outcomes at discharge were less likely to have pre-arrest hypotension (23 % vs 36 %, $P < 0.001$) or renal insufficiency (62 % vs 72 %, $P < 0.001$) and were less likely to be placed on mechanical ventilation (57 % vs 71 %, $P < 0.001$), have intra-arterial catheters (11 % vs 16 %, $P = 0.004$), or be administered vasoactive agents (24 % vs 41 %, $P < 0.001$) compared patients who died or survived with neurological deficits.

Supplementary Table 2 assesses the differences between the non-COVID-19 derivation cohorts (used to derive the LR, XGB, and MLP models) and the COVID-19 validation cohorts that met our inclusion criteria. We note skewness in a few aspects: our COVID-19 cohort was less likely to be female (36 % vs 43 %, $P < 0.001$), more likely to be Black (28 % vs 22 %, $P < 0.001$), experienced a shorter duration of arrests (8 min vs 10 min, $P < 0.001$), more likely to be in an intensive care setting (70 % vs 60 %, $P < 0.001$), and more likely to have an AED used during arrest (45 % vs 23 %, $P < 0.001$), compared to our derivation non-COVID-19 cohort. In addition, there were also significant differences observed in initial rhythm, with 60 % of COVID-19 patients experiencing a pulseless electrical activity arrest compared to 45 % in the non-COVID-19 patient cohort ($P < 0.001$). Patients in our

Table 1 – Clinical and Arrest Characteristics of Resuscitated Patients with COVID-19 with and without Favorable Neurological Outcome at Discharge.

Variable type	Variable	Patients with favorable neurological outcome ($n = 484$)	Patients without favorable neurological outcome ($n = 3641$)	*P*-value
Demographics	Age, mean (sd)	60.7 (13.7)	65.4 (13.1)	<0.001
	Female sex, n (%)	187 (38.6%)	1311 (36.0%)	0.28
	Race, n (%)			
	Black	127 (26.2%)	1013 (27.8%)	0.336
	White	279 (57.7%)	1958 (53.8%)	
	Other	78 (16.1%)	646 (17.7%)	
	Missing	0 (0%)	24 (0.7%)	
Characteristics of Arrest, n (%)	Initial Cardiac Arrest Rhythm			
	Asystole	100 (20.7%)	828 (22.7%)	0.004
	Pulseless Electrical Activity	274 (56.6%)	2214 (60.8%)	
	VT/VF T2FS<2min	32 (6.6%)	181 (5.0%)	
	VT/VF T2FS 2-3	23 (4.8%)	72 (2.0%)	
	VT/VF T2FS 3-4	3 (0.6%)	11 (0.3%)	
	VT/VF T2FS 4-5	1 (0.2%)	9 (0.2%)	
	VT/VF T2FS >5min	6 (1.2%)	43 (1.2%)	
	Unknown	45 (9.3%)	283 (7.8%)	
	Duration of Resuscitation, minutes, median (IQR)	5 (3–10)	8 (4–16)	<0.001
	Hospital Location			
	Telemetry	100 (20.7%)	560 (15.4%)	0.009
	Intensive Care Unit	315 (65.1%)	2579 (70.8%)	
	Inpatient	69 (14.2%)	502 (13.8%)	
	Time and Day of Arrest			
	Night	132 (27.3%)	1096 (30.1%)	0.226
	Weekend	143 (29.5%)	1158 (31.8%)	0.341
	Use of AED			
	Yes	202 (41.7%)	1668 (45.8%)	0.002
	No	245 (50.6%)	1551 (42.6%)	
	Not used-by-facility/NA	37 (7.7%)	422 (11.6%)	
	CPC Score prior to arrest			<0.001

CPC: Cerebral Performance Score.
VT: Ventricular Tachycardia.
VF: Ventricular Fibrillation.
T2FS: Time to First Shock.
IQR: Interquartile Range.
AED: Automated External Defibrillator.

Table 2 – Pre-Existing Conditions and Pre-Arrest Interventions for COVID-19 Resuscitated Patients With and Without Favorable Neurological Outcome at Discharge.

Variable type	Variable	Patients with favorable neurological outcome (n = 484)	Patients without favorable neurological outcome (n = 3641)	P-value
Pre-Existing Conditions, n (%)	Acute CNS Non-Stroke Event	63 (13.0%)	538 (14.8%)	0.336
	Acute Stroke	17 (3.5%)	123 (3.4%)	0.984
	Baseline Depression in CNS function	27 (5.6%)	293 (8.0%)	0.0692
	HF this admission	42 (8.7%)	321 (8.8%)	0.987
	HF prior admission	88 (18.2%)	677 (18.6%)	0.875
	Diabetes Mellitus	217 (44.8%)	1737 (47.7%)	0.254
	Hepatic Insufficiency	36 (7.4)	354 (9.7)	0.126
	Hypotension	113 (23.3%)	1319 (36.2%)	<0.001
	Major Trauma	17 (3.5%)	113 (3.1%)	0.730
	Malignancy	25 (5.2%)	240 (6.6%)	0.270
	Metabolic or Electrolyte Abnormality	140 (28.9%)	1296 (35.6%)	0.004
	Myocardial Infarction This Admission	43 (8.9%)	283 (7.8%)	0.446
	Myocardial Infarction Prior to This Admissions	47 (9.7%)	405 (11.1%)	0.391
	Pneumonia	238 (49.2%)	2108 (57.9%)	<0.001
	Renal Insufficiency	162 (33.5%)	1613 (44.3%)	<0.001
	Respiratory Insufficiency	302 (62.4%)	2613 (71.8%)	<0.001
Interventions in Place Prior to Arrest, n (%)	Assisted or Mechanical Ventilation	275 (56.8%)	2568 (70.5%)	<0.001
	Intra-arterial Catheter	54 (11.2%)	594 (16.3%)	0.004
	ECG Monitor	425 (87.8%)	3245 (89.1%)	0.430
	Pulse Oximeter	407 (84.1%)	3110 (85.4%)	0.481
	Vasoactive Agent	115 (23.8%)	1489 (40.9%)	<0.001
	Dialysis	18 (3.7%)	195 (5.4%)	0.156
	Implantable Cardiac Defibrillator	7 (1.4%)	45 (1.2%)	0.863

CNS: Central Nervous System.
HF: Heart Failure.
ECG: Electrocardiogram.

COVID-19 validation dataset were also more likely to have been placed on a variety of pre-arrest interventions (see **Supplementary** Table 2) and have higher incidence of pre-existing conditions such as hypotension (35 % vs 26 %, $P < 0.001$), metabolic abnormalities (35 % vs 19 %, $P < 0.001$), pneumonia (57 % vs 15 %, $P < 0.001$), renal (43 % vs 37 %) and respiratory (71 % vs 44 %, $P < 0.001$) insufficiency, and diabetes (47 % vs 33 %, $P < 0.001$), than patients in our non-COVID-19 derivation cohort.

Model performance

Table 3 describes the final AUCs for all the models used in this study. All our prior models (LR, XGBoost, and MLP) derived using non-COVID-19 resuscitation survivors outperformed CASPRI in terms of discrimination for COVID-19 resuscitation survivors with a favorable neurological outcome. The XGBoost model outperformed the LR model (AUC 0.75 vs 0.73, $P < 0.001$), but was similar to the MLP (AUC 0.75 vs 0.74, $P = 0.724$) in discriminating patients with the outcome from the patients without. Notably, there was no significant improvement in discrimination after transfer learning in comparison to the original MLP model (AUC: 0.74 vs 0.74, $P = 0.779$) and the XGBoost model (AUC 0.74 vs 0.75, $P = 0.940$).

Fig. 1 depicts the calibration curves indicating agreement between predicted and actual probabilities of the outcome. A perfect model calibration line will have a slope of 1 and an intercept of 0, indicating complete agreement (dashed line), and will be associated with a low unreliability index.[16] We note that, at the thresholds specified by Chan et al.,[4] the CASPRI model does not match the true prevalence of favorable neurological outcomes in the cohort of COVID-19 patients who survived resuscitation. Among our prior models, the LR and XGBoost show good calibration (LR model U 0.01, intercept −0.52, slope 0.81, $P < 0.001$; XGboost model U 0.01, intercept −0.49, slope 0.87, $P < 0.001$) while the MLP model performed the worst (U 0.17, intercept −0.69 slope 0.37, $P < 0.001$). However, transfer learning improved the calibration of the neural network model to outperform all models (U 0.00, intercept = -0.24, slope 0.89, $P = 0.073$).

Supplementary Table 3 compares the sensitivity, specificity, positive, and negative predictive values for CASPRI, XGBoost, and MLP-Transfer models. Overall, the accuracy metrics were very similar between the XGBoost and the MLP-Transfer models. At a sensitivity of 81 %, the XGBoost model had higher specificity (52 % vs 40 %), higher positive predictive value (18 % vs 15 %), and a slightly higher negative predictive value (95 %, 94 %) in detecting patients with favorable neurological outcomes, in comparison to CASPRI. Further, at a similar specificity (69 % for CASPRI and 68 % for the XGBoost), the XGBoost had a higher sensitivity (69 % vs 55 %), higher positive predictive value (22 % vs 19 %), and higher negative predictive value (94 % vs 92 %) than the CASPRI model. With a 5 %

Table 3 – Comparison of Model Performances for Predicting Favorable Neurological Outcome at Discharge in Resuscitation Survivors with COVID.

Model	AUC, 95%CI	*P*-value*
CASPRI	0.67 (0.65–0.70)	–
LR	0.73 (0.71–0.75)	<0.001
MLP	0.74 (0.72–0.77)	<0.001
MLP with transfer learning	0.74 (0.72–0.76)	<0.001
XGBoost	0.75 (0.73–0.77)	<0.001

AUC: Area Under the receiver operating characteristic Curve.
CI: Confidence Interval.
CASPRI: Cardiac Arrest Survival Post-Resuscitation In-hospital score.
LR: Logistic Regression.
MLP: Multi-Layer Perceptron.
XGBoost: eXtreme Gradient Boosted machine.
*In comparison with CASPRI.

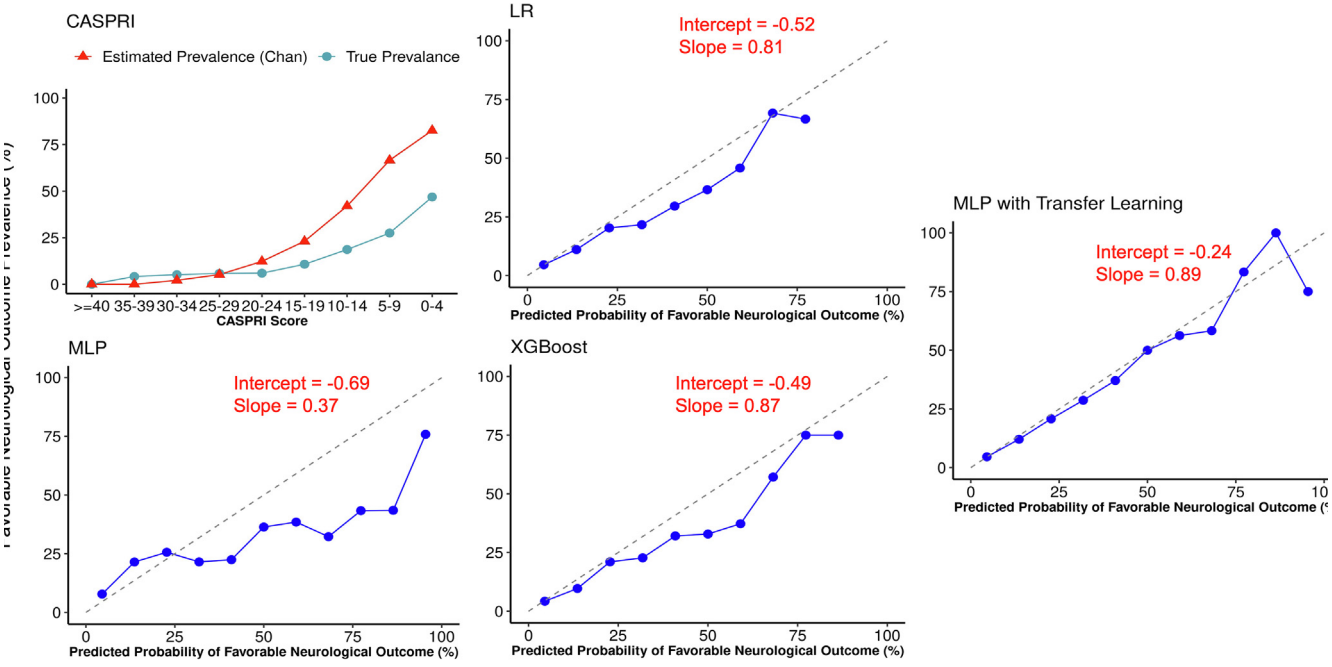

Fig. 1 – Calibration plots for CASPRI score (depicted on a reverse score scale) and the machine learning models demonstrating alignment between predicted probability of non-favorable neurological outcome at discharge against true outcome rate in COVID-19 resuscitation survivors.

or lower likelihood of surviving to discharge with favorable neurological outcomes (at XGBoost thresholds of $\leq$ 12 and Inverted CASPRI score $\leq$ 27), the XGBoost model had a higher sensitivity than CASPRI (24 % [95 %CI: 20 %-28 %] vs 19 % [95 %CI: 15 %–22 %]).

Fig. 2 depicts the variables most important for predicting the favorable neurological outcomes in the non-COVID-19 cohort for the XGBoost model and in the COVID-19 cohort for the MLP-Transfer model. Variables that were important for predicting outcomes in non-COVID 19 patients include admission CPC score, duration of resuscitation, initial cardiac rhythm, and age, consistent with our previous study.[5] These variables were similarly important for predicting neurological outcomes in the COVID-19 population. However, mechanical ventilation and pneumonia were also noted to be among the most important variables used by the MLP-Transfer model. **Supplementary Table 4** compares the discrimination and calibration performance of all machine learning models in the COVID-19 and the non-COVID-19 validation cohort from our prior study.[5] Missing value percentages for all predictors are shown in **Supplementary Table 5**.

Discussion

In this study, we compared the performance of existing models developed in patients without COVID-19 to predict favorable neurological outcomes in a population of more than 4,000 COVID-19 in-hospital cardiac arrest survivors from 241 hospitals from the GWTG-Resuscitation COVID-19 registry. Among our prior models, the gradient boosted machine outperformed CASPRI, a parsimonious score developed for easy scoring. The gradient boosted

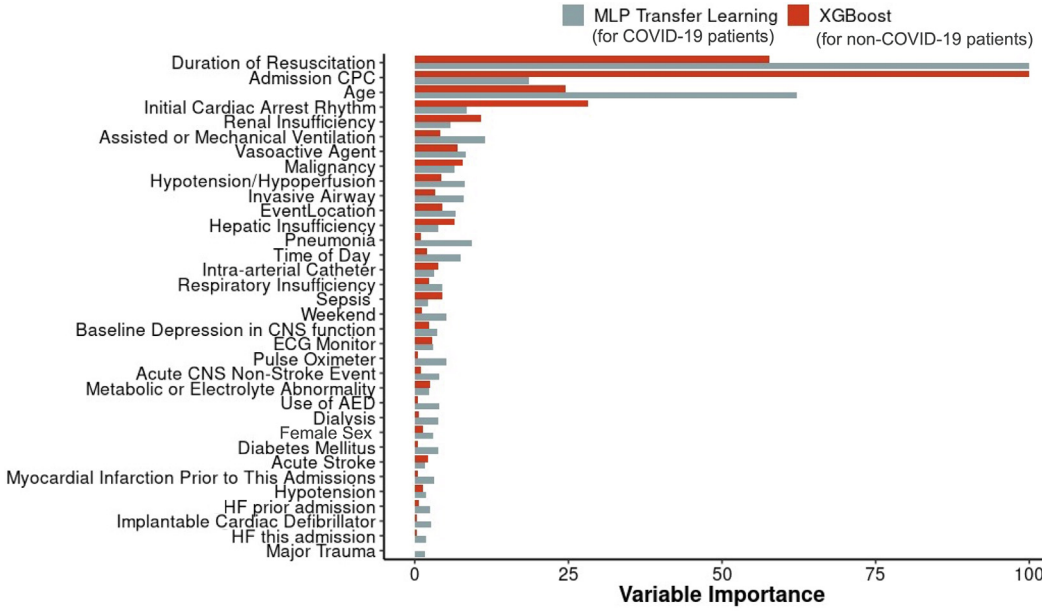

Fig. 2 – Importance of variables from the XGBoost model for predicting favorable neurological outcomes in non-COVID patients and the MLP transfer learning model for predicting favorable neurologic outcomes in COVID patients. Variable importance was calculated through permutation methods that measure dropout loss.

machine also outperformed other machine learning-based models in terms of discrimination, although all models overestimated the likelihood of survival with a favorable neurologic outcome in a substantial proportion of patients. We also demonstrated that transfer learning methods that adapted our neural network model to the COVID-19 population did not improve discrimination but did improve calibration performance. These results suggest that models developed in non-COVID-19 patients can discriminate well between those with and without favorable neurologic outcomes, but their predicted probabilities will be overly optimistic for many patients unless transfer learning is used.

Assessment of neurological status in survivors of cardiac arrest is difficult as these patients are often intubated, which is often distressing to families entrusted with goals-of-care decision making. These challenges are amplified in patients with COVID-19. First, rates of both overall post-arrest survival and survival to discharge without neurological deficits are low in COVID-19 patients who experience cardiac arrests.[6–11] In fact, in our study, we found that only 12 % of COVID-19 patients who survived their initial resuscitation were discharged alive with a favorable neurologic status, compared to 24 % of non-COVID-19 patients in our prior study. Second, prognostication in survivors is further complicated when clinical contact is recommended to be kept to a minimum, impeding proper physical and neurological assessment.[12,13] Therefore, the assessment of published models and tools specifically in COVID-19 patients is critical to improving prognostication after an in-hospital cardiac arrest.

Prior studies have utilized the large-scale, multicenter, national GWTG-R registry to develop scores such as the CASRPI and the Good Outcome Following Attempted Resuscitation (GO-FAR) scores for neurological prognostication.[4,17] In a recent study, we demonstrated that significant gains can be obtained using machine learning methods.[5] However, all our models were developed and validated in patients without COVID-19. This study is the first to test the performance of models that predict neurological outcomes in resuscitated patients specifically in a large, multicenter cohort of COVID-19 IHCA survivors. The gradient-boosted model remained the best-performing model for discriminating patients with favorable neurological outcomes from those without. Thus, our model could potentially be utilized for assessing neurological outcomes for risk adjustment and quality-based initiatives in the current pandemic when hospital surges are common and clinical contact is low. However, the model was overly optimistic for predicting favorable outcomes for some patients, as illustrated by its calibration curve.

In our previous study, the machine learning models had marginal improvements in discrimination over the logistic regression when tested on non-COVID 19 patients (c-statistic 0.81 vs 0.79, $P < 0.001$). However, in this study, we note the superior performance of the gradient-boosted model over the logistic regression model in COVID-19 patients, suggesting that these models may be more generalizable than standard regression methods in this clinical situation. Further, our prior neural network demonstrated similar discrimination but worse calibration than the gradient-boosted method for this study. Calibration improved considerably when the neural network was trained to adapt to COVID-19 patients using transfer learning. Model calibration may be of high importance for cardiac arrest prognostication, especially if clinical decisions are tied to specific likelihoods (e.g., <5% predicted probability of a favorable outcome). Our models can be deployed within a hospital setting with varying technical resource requirements. Regression-based predictions are simple aggregations of multiplicative products between coefficients and feature values and are directly implementable in some EHR systems. Deploying the more accurate gradient-boosted machine behind an electronic health record interface requires embedding the model within a predictive model markup language (PMML) system or other similar infrastructure. Deployment of the transfer

learning-based MLP requires additional resources for computation and libraries for deep learning.

A comparison of global variable importance between the gradient boosted model and the neural network transfer learning model suggested commonalities as well as differences in variables important for accurate prediction of neurological status between non-COVID-19 and COVID-19 patients. Duration of resuscitation, initial cardiac arrest rhythm, admission CPC, and patient age were important to determining the likelihood of neurological survival in resuscitation survivors across both populations.[18–20] Mechanical ventilation and pneumonia had a higher variable importance in our transfer learning model that was trained on COVID-19 patients. Clinical manifestations of pneumonia in COVID-19 patients are dominant,[21] while the need for mechanical ventilation has traditionally been associated with poor survival in COVID-19 patients.

Our study has several limitations. First, we utilized retrospective data elements that are available within the GWTG-R registry and thus may be missing important predictors of neurological outcomes among COVID-19 patients. Second, our cohort included both confirmed and suspected COVID-19 cases, which may increase the heterogeneity of the results. Further validation is required to assess differences between patients suspected of COVID-19 in comparison to those confirmed with COVID-19. Other cohort selection biases may also exist based on return of spontaneous circulation criteria that may not be missing at random or by eliminating patients due to early withdrawal of treatment. Additionally, about 2 % of our excluded population had missing discharge CPC scores. While this is less than our original study on non-COVID-19 patients (about 5 % of exclusions had missing discharge CPC scores), it is potentially another source of selection bias. These limitations collectively underline the need for prospective validation prior to use in clinical practice. Third, while our model includes pre-existing conditions, our data does not further expand on details regarding IHCA etiology. Fourth, similar to published methods, our model cannot be used for assessing neurological outcomes in out-of-hospital cardiac arrests. We also stress that AUC is a global metric for assessing the performance of prediction models. Individual decisions regarding clinical care must be made at cut points that are set with clinically actionable responses in mind. Similar to our original non-COVID-19 model, our model is best suited for estimating chances of adequate neurological recovery for COVID-19 IHCA survivors for quality initiatives, risk-adjustment, or goal-of-care discussions after additional prospective validation. Finally, our study does not account for hospital factors or local pandemic responses that could impact assessment and resuscitation practices.

Conclusion

We validated the performance of published machine learning models for detecting resuscitation survivors with COVID-19 who are likely to be discharged with a favorable neurologic status. Our results highlight the utility of these models for predicting neurological outcomes in COVID-19 cardiac arrest survivors and the ability of transfer learning to improve model calibration.

CRediT authorship contribution statement

Anoop Mayampurath: Methodology, Software, Validation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Fereshteh Bashiri:** Methodology, Software, Writing – review & editing. **Raffi Hagopian:** Methodology, Software, Writing – review & editing. **Laura Venable:** Methodology, Software, Writing – review & editing. **Kyle Carey:** Data curation, Software, Writing – review & editing. **Dana Edelson:** Conceptualization, Investigation, Writing – review & editing. **Matthew Churpek:** Conceptualization, Methodology, Investigation, Formal analysis, Supervision, Writing – original draft, Writing – review & editing.

Acknowledgements

We thank Mary Akel and Madeline Oguss for administrative assistance. We also thank the members of the American Heart Association's Get With The Guidelines Adult Research Task Force: Anne Grossestreuer PhD; Ari Moskowitz MD; Dana Edelson MD MS; Joseph Ornato MD; Mary Ann Peberdy MD; Matthew Churpek MD MPH PhD; Monique Anderson Starks MD MHS; Paul Chan MD MSc; Saket Girotra MBBS SM; Sarah Perman MD MSCE; Zachary Goldberger MD MS. IQVIA (Parsippany, New Jersey) serves as the data collection (through their Patient Management Tool – PMT™) and coordination center for GWTG. The University of Pennsylvania serves as the data analytic center and has an agreement to prepare the data for research purposes. The Get With The Guidelines®-Resuscitation program is provided by the American Heart Association. The American Heart Association Precision Medicine Platform (https://precision.heart.org/) was used for data analysis. All participating institutions were required to comply with local regulatory and privacy guidelines and, if required, to secure institutional review board approval. Because data were used primarily at the local site for quality improvement, sites were granted a waiver of informed consent under the common rule. Dr. Mayampurath is supported by a career development award from the National Heart, Lung, and Blood Institute (K01HL148390). Dr. Churpek is supported by R01s from NIGMS (R01 GM123193), NHLBI (R01 HL157262), NIDDK (R01-DK126933), and a PRMRP grant from DOD (W81XWH-21-1-0009).

Conflicts and Disclosures

Dr. Mayampurath is supported by a career development award from the National Heart, Lung, and Blood Institute (K01HL148390). Dr. Churpek is supported by R01s from NIGMS (R01 GM123193), NHLBI (R01 HL157262), NIDDK (R01-DK126933), and a PRMRP grant from DOD (W81XWH-21-1-0009). Drs. Churpek and Edelson have a patent pending (ARCD. P0535US.P2) for risk stratification algorithms for hospitalized patients and has received research support from EarlySense (Tel Aviv, Israel). Dr. Edelson has received research support and honoraria from Philips Healthcare (Andover, MA). Dr. Edelson has ownership interest in AgileMD (San Francisco, CA), which licenses eCART, a patient risk analytic.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resuscitation.2022.07.018.

Author details

for the American Heart Association's Get With The Guidelines®-Resuscitation Investigators [a]*Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, WI, United States* [b]*Department of Medicine, University of Wisconsin, Madison, WI, United States* [c]*Department of Medicine, Weill Cornell Medicine, New York, NY, United States* [d]*Department of Medicine, University of Chicago, Chicago, IL, United States*

REFERENCES

1. Ebell MH, Becker LA, Barry HC, Hagen M. Survival After In-Hospital Cardiopulmonary Resuscitation. J Gen Intern Med 1998;13:805–16. https://doi.org/10.1046/j.1525-1497.1998.00244.x.
2. Girotra S, Nallamothu BK, Spertus JA, Li Y, Krumholz HM, Chan PS. Trends in Survival after In-Hospital Cardiac Arrest. N Engl J Med 2012;367:1912–20. https://doi.org/10.1056/NEJMoa1109148.
3. Geocadin RG, Callaway CW, Fink EL, et al. Standards for Studies of Neurological Prognostication in Comatose Survivors of Cardiac Arrest: A Scientific Statement From the American Heart Association. Circulation 2019;140:e517–42. https://doi.org/10.1161/CIR.0000000000000702.
4. Chan PS, Spertus JA, Krumholz HM, et al. A Validated Prediction Tool for Initial Survivors of In-Hospital Cardiac Arrest. Arch Intern Med 2012;172:947–53. https://doi.org/10.1001/archinternmed.2012.2050.
5. Mayampurath A, Hagopian R, Venable L, et al. Comparison of Machine Learning Methods for Predicting Outcomes After In-Hospital Cardiac Arrest. Crit Care Med 2021. https://doi.org/10.1097/CCM.0000000000005286.
6. Mitchell OJL, Yuriditsky E, Johnson NJ, et al. In-hospital cardiac arrest in patients with coronavirus 2019. Resuscitation 2021;160:72–8. https://doi.org/10.1016/j.resuscitation.2021.01.012.
7. Hayek SS, Brenner SK, Azam TU, et al. In-hospital cardiac arrest in critically ill patients with covid-19: multicenter cohort study. BMJ 2020;371. https://doi.org/10.1136/bmj.m3513 m3513.
8. Sultanian P, Lundgren P, Strömsöe A, et al. Cardiac arrest in COVID-19: characteristics and outcomes of in- and out-of-hospital cardiac arrest. A report from the Swedish Registry for Cardiopulmonary Resuscitation. Eur Heart J 2021;42:1094–106. https://doi.org/10.1093/eurheartj/ehaa1067.
9. Acharya P, Ranka S, Sethi P, et al. Incidence, Predictors, and Outcomes of In-Hospital Cardiac Arrest in COVID-19 Patients Admitted to Intensive and Non-Intensive Care Units: Insights From the AHA COVID-19 CVD Registry. J Am Heart Assoc 2021;10. https://doi.org/10.1161/JAHA.120.021204 e021204.
10. Thapa SB, Kakar TS, Mayer C, Khanal D. Clinical Outcomes of In-Hospital Cardiac Arrest in COVID-19. JAMA Internal Med 2021;181:279–81. https://doi.org/10.1001/jamainternmed.2020.4796.
11. Bhardwaj A, Alwakeel M, Saleem T, et al. A Multicenter Evaluation of Survival After In-Hospital Cardiac Arrest in Coronavirus Disease 2019 Patients. Crit Care Explor 2021;3. https://doi.org/10.1097/CCE.0000000000000425 e0425.
12. Edelson DP, Sasson C, Chan PS, et al. Interim Guidance for Basic and Advanced Life Support in Adults, Children, and Neonates With Suspected or Confirmed COVID-19. Circulation 2020;141:e933–43. https://doi.org/10.1161/CIRCULATIONAHA.120.047463.
13. Ranney ML, Griffeth V, Jha AK. Critical Supply Shortages - The Need for Ventilators and Personal Protective Equipment during the Covid-19 Pandemic. N Engl J Med 2020;382. https://doi.org/10.1056/NEJMp2006141 e41.
14. Holder AL, Shashikumar SP, Wardi G, Buchman TG, Nemati S. A Locally Optimized Data-Driven Tool to Predict Sepsis-Associated Vasopressor Use in the ICU. Crit Care Med 2021;49:e1196–205. https://doi.org/10.1097/CCM.0000000000005175.
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.
16. Harrell FE. Binary Logistic Regression. In: Harrell Jr, Frank E, editors. Springer Series in Statistics. Springer International Publishing; 2015. p. 219–74. doi:10.1007/978-3-319-19425-7_10.
17. Ebell MH, Jang W, Shen Y, Geocadin RG. Get With the Guidelines-Resuscitation Investigators. Development and validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score to predict neurologically intact survival after in-hospital cardiopulmonary resuscitation. JAMA Int Med 2013;173:1872–8. https://doi.org/10.1001/jamainternmed.2013.10037.
18. Andersen LW, Holmberg MJ, Berg KM, Donnino MW, Granfeldt A. In-Hospital Cardiac Arrest: A Review. JAMA 2019;321:1200–10. https://doi.org/10.1001/jama.2019.1696.
19. Coppler PJ, Elmer J, Calderon L, et al. Validation of the Pittsburgh Cardiac Arrest Category illness severity score. Resuscitation 2015;89:86–92. https://doi.org/10.1016/j.resuscitation.2015.01.020.
20. Rohlin O, Taeri T, Netzereab S, Ullemark E, Djärv T. Duration of CPR and impact on 30-day survival after ROSC for in-hospital cardiac arrest-A Swedish cohort study. Resuscitation 2018;132:1–5. https://doi.org/10.1016/j.resuscitation.2018.08.017.
21. Attaway AH, Scheraga RG, Bhimraj A, Biehl M, Hatipoğlu U. Severe covid-19 pneumonia: pathogenesis and clinical management. BMJ 2021;372. https://doi.org/10.1136/bmj.n436 n436.