Machine learning predicting mortality in sarcoidosis patients admitted for acute heart failure

Qiying Dai, MD,* Akil A. Sherif, MD,† Chengyue Jin, MD,‡ Yongbin Chen, MD, PhD,§ Peng Cai, MS,¶ Pengyang Li, MD‖

*From the *Division of Cardiology, Mayo Clinic, Rochester, Minnesota, †Division of Cardiology, Saint Vincent Hospital, Worcester, Massachusetts, ‡Division of Cardiology, Mount Sinai Beth Israel Medical Center, New York, New York, §Biochemistry and Molecular Biology, Mayo Clinic, Rochester, Minnesota, ¶Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, Massachusetts, and ‖Division of Cardiology, Pauley Heart Center, Virginia Commonwealth University, Richmond, Virginia.*

BACKGROUND Sarcoidosis with cardiac involvement, although rare, has a worse prognosis than sarcoidosis involving other organ systems.

OBJECTIVE We used a large dataset to train machine learning models to predict in-hospital mortality among sarcoidosis patients admitted with heart failure (HF).

METHOD Utilizing the National Inpatient Sample, we identified 4659 patients hospitalized with a primary diagnosis of HF. In this cohort, we identified patients with a secondary diagnosis of sarcoidosis using *International Statistical Classification of Disease, Tenth Revision* (ICD-10) codes. Patients were separated into a training group and a testing group in a 7:3 ratio. Least absolute shrinkage and selection operator regression was used to select variables to prevent model overfitting or underfitting. For machine learning models, logistic regression, random forest, and XGBoosting were applied in the training group. Parameters in each of the models were tuned using the GridSearchCV function. After training, all models were further validated in the testing group. Models were then evaluated using the area under curve (AUC) score, sensitivity, and specificity.

RESULTS A total of 2.3% of sarcoidosis patients died in HF admission. Our machine learning model analysis found the RF model to have the highest AUC score and sensitivity. Feature analysis found that comorbid arrhythmias and fluid electrolyte disorders were the strongest factors in predicting in-hospital mortality.

CONCLUSION Machine learning methods can be useful in identifying predictors of in-hospital mortality in a given dataset.

KEYWORDS Heart failure; In-hospital mortality; Machine learning; National Inpatient Sample; Sarcoidosis

(Cardiovascular Digital Health Journal 2022;3:297–304)

Introduction

Sarcoidosis is an autoimmune granulomatous disease that frequently involves multiple organ systems. In the past, cardiac involvement in sarcoidosis patients was thought to be rare. However, with improvements in image technology, the detection rate of cardiac involvement was reported to be as high as 26%.[1] About 20%–30% of patients with sarcoidosis have cardiac involvement.[2] Clinical presentations of cardiac sarcoidosis (CS) vary and can present with atrioventricular block, ventricular arrhythmia, cardiomyopathy, heart failure (HF), pericardial disease, or sudden cardiac death.[3]

The presence of sarcoidosis increases the risk of developing HF. In a study following 12,042 sarcoidosis patients for 8.2 years, it was estimated that, among patients with sarcoidosis, the absolute 10-year risk of developing HF was 3.18%, which is 3 times higher than that of the general population.[4] HF is also one of the most important predictors of mortality in patients with sarcoidosis.[5] Progressive HF subsequent to sarcoidosis has been shown to be accountable for 25% of deaths in patients with CS. This makes it the second most common cause of death after sudden death in these patients.[6] In a study of 95 Japanese patients with CS, it was shown that 73% of the patients died of congestive HF. In addition, the study also demonstrated increased mortality, with a hazard ratio of 7.72 per New York Heart Association functional class increase.[7] A multivariate regression analysis from another study involving 351 CS patients with follow-up times ranging from 6 months to 29.7 years found HF to be an independent predictor of mortality, especially among those with ejection fraction <35%.[8] HF caused by CS is thought to occur due to infiltration of noncaseating granulomas,

Address reprint requests and correspondence: Dr Pengyang Li, Division of Cardiology, Pauley Heart Center, Virginia Commonwealth University, Richmond, VA 23219. E-mail address: leelpy0109@gmail.com.

2666-6936/

https://doi.org/10.1016/j.cvdhj.2022.08.001

KEY FINDINGS

- Machine learning algorithms are useful tools in identifying predictors of in-hospital mortality in sarcoidosis patients admitted for heart failure.
- In using clinical characteristics as variables to build various machine learning models, we found the random forest model to have the best performance.
- Future studies involving more detailed clinical information including laboratory parameters, imaging, and medication data are required to improve the model performance.

regional metabolic abnormalities, or microvascular vasoconstriction. These pathophysiological changes are detected by imaging studies, including cardiac magnetic resonance, dipyridamole thallium-201 myocardial scan, and technetium-99m sestamibi tomoscintigraphy.[9–11]

Given the relationship between HF and sarcoidosis and its implication on patient mortality, we believed it is essential to evaluate early estimators of in-hospital mortality in this group of patients. Machine learning (ML) methods enable us to rapidly generate prediction models based on thousands of clinical patterns.[12] This is achieved by using different algorithms to learn from a dataset by identifying patterns in the given data. In recent years, ML techniques have been increasingly utilized in cardiology clinical research.[13–16] We used 3 different ML algorithms to explore the feasibility and accuracy of predicting in-hospital mortality in sarcoidosis patients admitted for HF.

Methods

Data source

The data used for analysis in our study were obtained from the National Inpatient Sample (NIS) database. The NIS is part of an extensive database developed for the Healthcare Cost and Utilization Project. It involves more than 7 million hospital stays each year and has been the largest publicly available in-patient health care dataset. Individual patient information in the NIS databases is de-identified, so institutional review board approval is not required in studies utilizing this dataset.[17] Starting in 2016, all diagnoses included in the NIS database were reported using the *International Classification of Diseases, Tenth Revision, Clinical Modification* (ICD-10-CM) coding system. All the variables collected in this study were identified using ICD-10 codes.

We selected all patients with a primary diagnosis of HF from January 1, 2016, to December 31, 2019. Patients with comorbid sarcoidosis as a secondary diagnosis were then identified using ICD-10-CM codes for sarcoidosis. The endpoint was defined as in-hospital mortality. Individuals with missing data on age, gender, race, or in-hospital mortality information were excluded. The variables used in this study were based on the Elixhauser comorbidity scores.[18] ICD-10-CM codes of these comorbidities are given in Supplemental Table 1. Categorical variables were expressed using dummy variables for purposes of algorithm training.[19]

Study design

The data processing and ML algorithms were developed in Python Version 3.8.8. First, patients were divided into a training group (3261 patients; 81 [2.5%] died; 3180 [97.5%] did not) and a testing group (1398 patients; 25 [1.9%] died; 1373 [98.1%] did not) in a 7:3 ratio.[16] Because 2.3% patients died during hospitalization, which is considered an imbalanced database, we used the oversampling method to achieve an equal amount of patients with in-hospital mortality.[20] This translates to a 1:1 ratio of patients who died in-hospital vs patients who did not from the 81:3180 ratio. The total number of patients involved after oversampling was 6360. The process of dividing the groups and oversampling is shown in Figure 1.

Before developing the algorithms, the nullity correlation of every variable was determined (Figure 2). Nullity correlation refers to the relationship between 2 quantities. It ranges from –1 to 1. A value of "0" indicates no association between 2 variables; "1" indicates that both variables appear at the same time, and "–1" denotes that 1 variable appears while the other definitely does not. To prevent model overfitting, least absolute shrinkage and selection operator (LASSO) regression was used to select variables that best contribute to HF admissions.[21,22]

Statistical analysis

Categorical variables and binary variables were compared using χ^2 tests. Continuous variables were compared using the Student t test. Statistical analysis was performed using Python Version 3.8.8.

ML algorithm

We used 3 classification-based ML models: logistic regression (LR), random forest (RF), and XGBoosting to predict in-hospital mortality. LR, also referred to as sigmoid function, combines variables linearly using weights or coefficient values to predict an outcome.[23] In this model, the parameter we tuned is "C", which equals 1/lambda. Parameter "C" regulates the complexity of the model to prevent overfitting/underfitting. We used GridSearchCV, a hyperparameter tuning tool from the sklearn python package, to exhaustively search from 10^{-4} to 10^{4} to find "C" representative of the best predictive value. RF is a nonlinear algorithm that ensembles bunches of classification models (decision tree models) to boost the accuracy of prediction; that is, it uses multiple models to make the exact prediction. A system called majority voting is used to decide the final outcome after receiving the greatest number of votes from each classification model.[24] In this model, we tuned the parameters n_estimators, max_deth, max_features, min_samples_split,

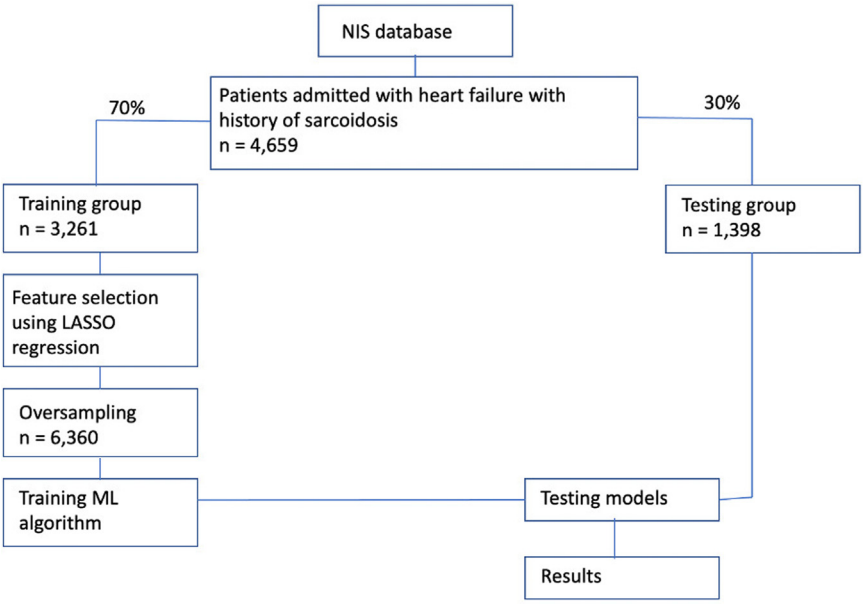

Figure 1 Flowchart of machine learning. LASSO = least absolute shrinkage and selection operator; ML = machine learning; NIS = National Inpatient Sample.

and min_samples_leaf by running RandomizedSearchCV, another hyperparameter tuning tool available in the sklearn python package. We used RandomizedSearchCV to run through a large scale of numbers to find an approximate best match. We then used GridSearchCV to exhaustively search through all the numbers around the approximate best match from RandomSearchCV. XGBoosting, in contrast, implements gradient boosted decision trees to predict results. Compared to RF, XGBoosting makes predictions sequentially rather than dependently. That is, instead of voting based on results from multiple decision trees, XGBoosting leverages the patterns in residuals and strengthens the model with weak predictions.[25] Parameters tuned in this model include n_estimators, min__child_weight, max_depth, gamma, subsample, and colsample_bytree. Similar to the RF model, we used RandomizedSearchCV to find the approximate best match for each parameter. We then used GridSearchCV to test each number around the approximate best match to find the best predictive value. All parameters were tuned with a goal of obtaining the maximum value for area under the receiver operating characteristics curve (AUC).[26] After that, AUC of each algorithm was presented with 95% confidence interval (CI). Lastly, feature importance of each variable in the algorithms was calculated respectively, along with the sensitivity and specificity of each algorithm.

Results:

Study population and baseline characteristics

A total of 4659 patients were included in this study. Among these patients, 106 (2.3%) died during HF hospitalization. Demographic features and comorbidities are given in Table 1. In terms of demographic features, patients who died during HF hospitalization seem to be older than those who did not (68.60 ± 12.85 years vs 63.36 ± 13.04; $P < .001$). No significant differences in gender (53.8% vs 57.9%; $P = .45$) or race were noted. Regarding comorbidities, patients who died during HF hospitalization were found to have higher rates of cardiac arrhythmia (71.7% vs 49.9%; $P < .001$), renal failure (64.2% vs 53.6%; $P = .041$), liver disease (18.9% vs 6.7%; $P < .001$), coagulopathy (17.0% vs 6.9%; $P < .001$), and weight loss (13.2% vs 5.4%; $P = .001$) than those who did not. No significant differences were noted in the distributions of CS (7.5% vs 6.2%; $P = .706$), congestive HF (83.0% vs 86.0%; $P = .471$), valvular disease (17.0% vs 19.5%; $P = .592$), pulmonary circulation disorders (44.3% vs 37.6%; $P = .191$), peripheral vascular disease (3.8% vs 4.4%; $P = .946$), uncomplicated hypertension (5.7% vs 7.4%; $P = .629$), complicated hypertension (15.1% vs 10.2%; $P = .137$), presence of paralysis (1.9% vs 0.4%; $P = .134$), neurological disorders (5.7% vs 3.1%; $P = .213$), chronic pulmonary disease (46.2% vs 49.5%; $P = .569$), uncomplicated and complicated diabetes ($P = .565$ and $P = 1$, respectively), hypothyroidism (16.0% vs 15.8%; $P = 1$), peptic ulcer disease (0.9% vs 0.5%; $P = 1$), acquired immunodeficiency syndrome (0.9% vs 0.2%; $P = .45$), lymphoma (0.9% vs 0.7%; $P = 1$), metastatic cancer (0.0% vs 0.6%; $P = .882$), solid tumors (0.9% vs 1.9%; $P = .728$), and rheumatoid disease (7.5% vs 5.6%; $P = .512$) between patients who died and those who did not.

ML models: Performance

AUC curves of the 3 trained ML models are shown in Figure 3. RF was found to have the highest AUC value

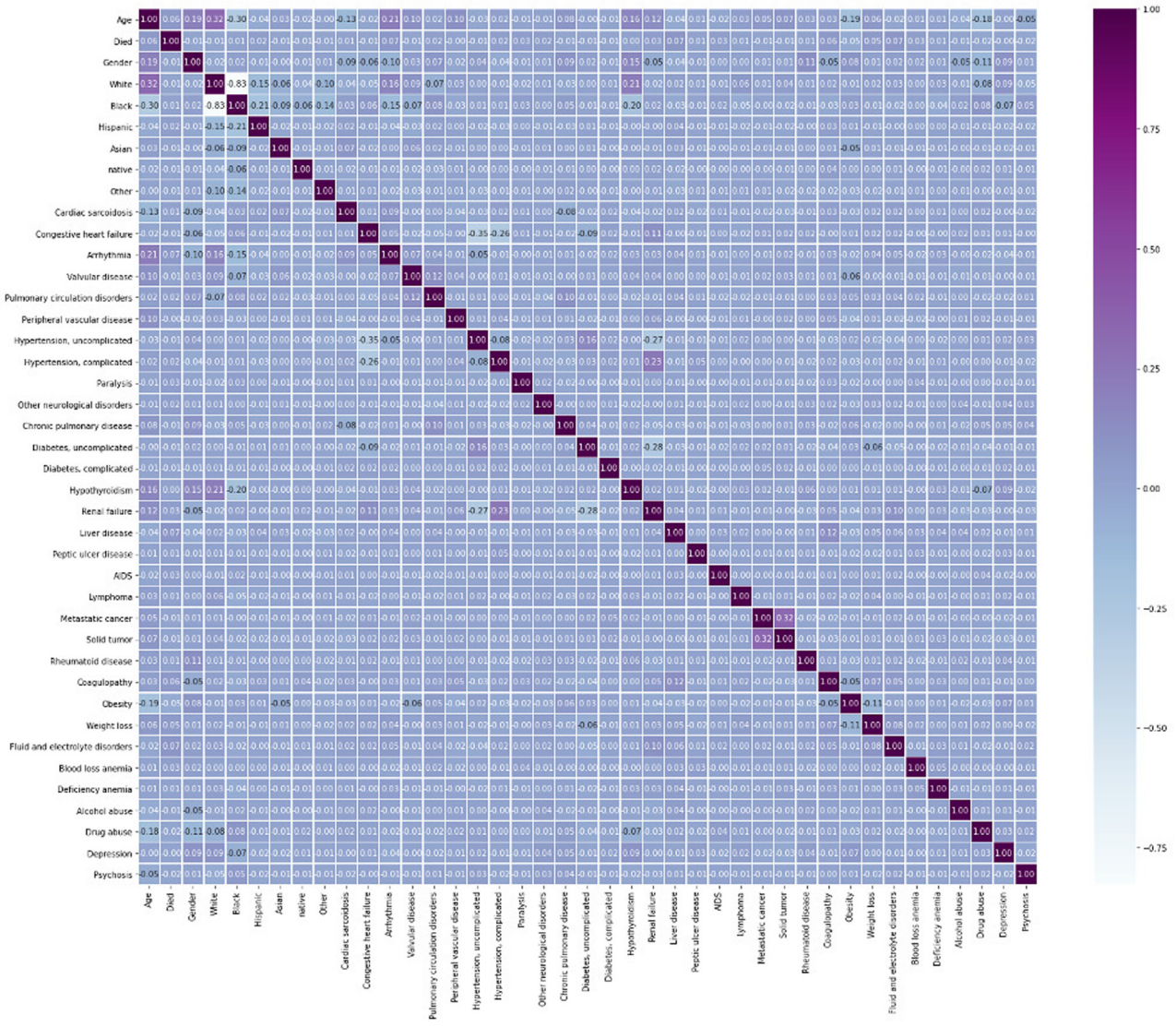

Figure 2 Nullity correlation of every variable.

(0.71; 95% CI 0.59–0.82). XGBoosting also showed a good AUC for predicting in-hospital mortality (0.70; 95% CI 0.58–0.81). The LR model algorithm showed a relatively poorer predicting capability (0.65; 95% CI 0.53–0.76) compared to the other 2 models. An evaluation of the 3 models is given in Table 2. RF was found to have the highest sensitivity (60.0%), whereas XGBoosting had the highest specificity (97.2%).

Feature importance from each trained model after tuning parameters is shown in Figure 4. In the LR model, paralysis was the strongest predictor of in-hospital mortality. In the RF model, fluid-electrolyte disorders seemed to be the strongest predictor of mortality, followed by age, cardiac arrhythmias, and liver disease. In the XGBoosting model, coagulopathy was the strongest predictor, followed by fluid-electrolyte disorders and cardiac arrhythmias.

Discussion

We developed 3 ML models to predict in-hospital mortality among sarcoidosis patients hospitalized for HF based on their clinical features. RF performed better than the other 2 models. The AUC score of the RF model in our study is 0.71, which is considered to be good.[27] Furthermore, feature importance from the RF model found fluid-electrolyte disorders, age, and cardiac arrhythmias to be the 3 most important factors contributing to in-hospital mortality. To our knowledge, we are the first group to analyze an ML-based algorithmic approach to predict in-hospital mortality among sarcoidosis patients hospitalized with HF.

Sarcoidosis is a granulomatous, infiltrative disease that can involve the myocardium, resulting in HF among other cardiac sequelae. With the advent of rhythm management with pacemakers and implantable defibrillators, the most

Table 1 Baseline features

	Not died	Died	*P* value
N	4553	106	
Age (y)	63.36 ± 13.04	68.60 ± 12.85	<.001
Gender (female)	2637 (57.9)	57 (53.8)	.45
Race			
White	1652 (36.3)	35 (33.0)	.556
Black	2489 (54.7)	60 (56.6)	.766
Hispanic	161 (3.5)	7 (6.6)	.158
Native American	16 (0.4)	0 (0.0)	1
Other	75 (1.6)	1 (0.9)	.859
Cardiac sarcoidosis	281 (6.2)	8 (7.5)	.706
Congestive heart failure	3914 (86.0)	88 (83.0)	.471
Cardiac arrhythmia	2271 (49.9)	76 (71.7)	<.001
Valvular disease	890 (19.5)	18 (17.0)	.592
Pulmonary circulation disorder	1713 (37.6)	47 (44.3)	.191
Peripheral vascular disease	200 (4.4)	4 (3.8)	.946
Hypertension, uncomplicated	336 (7.4)	6 (5.7)	.629
Hypertension, complicated	463 (10.2)	16 (15.1)	.137
Paralysis	19 (0.4)	2 (1.9)	.134
Neurological disorders	139 (3.1)	6 (5.7)	.213
Chronic pulmonary disease	2254 (49.5)	49 (46.2)	.569
Diabetes, uncomplicated	717 (15.7)	14 (13.2)	.565
Diabetes, complicated	13 (0.3)	0 (0.0)	1
Hypothyroidism	721 (15.8)	17 (16.0)	1
Renal failure	2442 (53.6)	68 (64.2)	.041
Liver disease	303 (6.7)	20 (18.9)	<.001
Peptic ulcer disease	23 (0.5)	1 (0.9)	1
AIDS	7 (0.2)	1 (0.9)	.45
Lymphoma	30 (0.7)	1 (0.9)	1
Metastatic cancer	27 (0.6)	0 (0.0)	.882
Solid tumor	86 (1.9)	1 (0.9)	.728
Rheumatoid disease	254 (5.6)	8 (7.5)	.512
Coagulopathy	316 (6.9)	18 (17.0)	<.001
Obesity	1617 (35.5)	22 (20.8)	.002
Weight loss	245 (5.4)	14 (13.2)	.001
Fluid and electrolyte disorders	1695 (37.2)	65 (61.3)	<.001
Blood loss anemia	25 (0.5)	2 (1.9)	.252
Deficiency anemia	51 (1.1)	2 (1.9)	.785
Alcohol abuse	12 (0.3)	0 (0.0)	1
Drug abuse	219 (4.8)	2 (1.9)	.243
Depression	596 (13.1)	13 (12.3)	.917
Psychosis	49 (1.1)	0 (0.0)	.554

Values are given as mean ± SD or n (%) unless otherwise indicated.
AIDS = acquired immunodeficiency syndrome.

common cause of death in these patients has shifted from sudden cardiac death to HF.[28] However, diagnosis of CS remains a challenge because about 50% of sarcoidosis patients are asymptomatic at the time of sarcoidosis diagnosis.[9] Additionally, CS causes patchy infiltration of the myocardium, which lowers the sensitivity of current diagnostic methods.[29] This leads to a significant number of undiagnosed CS cases. As a result, data regarding prevalence of CS among sarcoidosis patients with HF are lacking.

Whether presence of HF indicates progression to a late stage of sarcoidosis is unclear, as the presentation depends on the location and amount of myocardial involvement.[30] HF at the time of presentation is an independent predictor of survival.[31] Early detection of myocardial involvement and prompt treatment are associated with reduced mortality.[7,31–33] Little is known about the risk of mortality associated with clinical features among sarcoidosis patients with HF. Previous studies focused on sarcoidosis cardiomyopathy showed an overall in-hospital mortality of about 2.5%. In the study, age, peripheral vascular disease, chronic lung disease, liver disease, renal disease, and arrhythmias such as atrial fibrillation and ventricular fibrillation all were shown to contribute to mortality independently.[34] In our study, we established ML models to predict in-hospital mortality among sarcoidosis patients hospitalized with HF, based on their clinical characteristics.

We also identified comorbidities that significantly impacted outcome prediction by the ML algorithm. In RF, the feature importance plot showed fluid-electrolyte disorder to be the most important factor, followed by age, arrhythmias, and liver disease. This is consistent with the conclusions drawn by previous studies. Electrolyte and fluid imbalance is a common problem encountered in the management of HF in general, and failure to adequately address this is associated with poor clinical outcomes. Electrolyte disturbances such as hyponatremia,[35] hyperkalemia,[36] and fluid imbalance[37] are known to be closely associated with short-term mortality in patients with HF. A study of 73 patients with CS showed that age is a significant predictor of mortality.[38] A study based on the NIS database of 369,285 sarcoidosis-related hospitalizations showed atrial fibrillation to be the most common cardiac arrhythmia, followed by ventricular tachycardia. Individuals with arrhythmias were found to have higher in-hospital mortality.[39] Another study on 113 patients found arrhythmias to be the terminal incidence of 67% of CS-related deaths.[40] Sarcoidosis also causes liver disease. In a retrospective study of 286 sarcoidosis patients, 9.4% were found to have liver sarcoidosis, and 37% among them had significant clinical features including cirrhosis and portal hypertension.[41] Although it is unclear whether sarcoidosis can exacerbate pre-existing liver disease caused by other etiologies, no clear evidence suggesting associations of mortality was found in CS patients with concomitant liver disease.

One of the strengths of our study is the large sample size, which was achieved using the NIS database and resulted in substantial statistical power.[42–44] Larger samples of patients enable us to construct stronger ML algorithms. We found RF to be a good model, with high AUC scores for HF hospitalization prediction. RF had the highest sensitivity, whereas XGBoosting had the highest specificity. This finding highlights the applicability of ML technology in cardiovascular medicine.

Study limitations

Limitations associated with the use of the NIS database include lack of clinical details such as medication use,

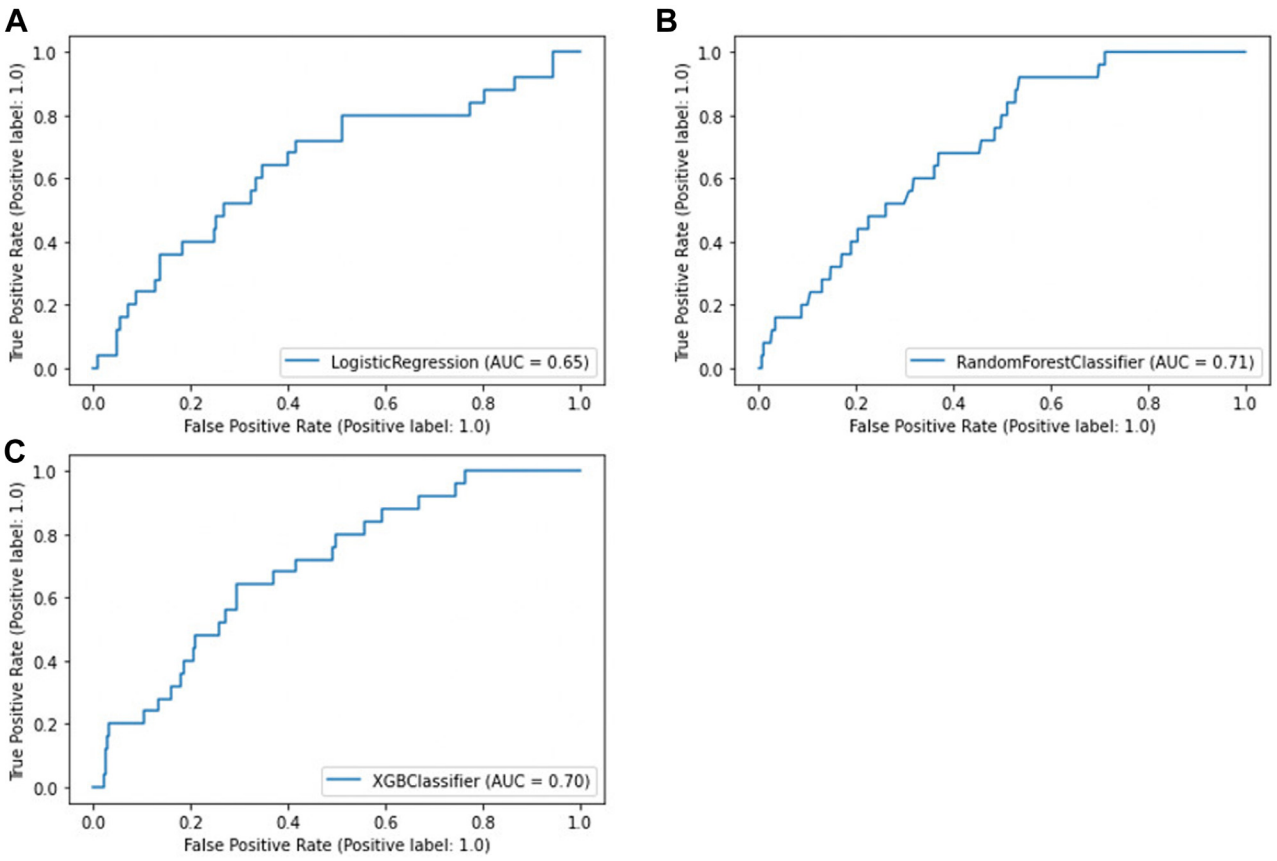

Figure 3 Presentation of AUC of 3 trained machine learning models. **A:** Logistic regression model. **B:** Random forest model. **C:** XGBoosting model. AUC = area under the receiver operating characteristic curve.

imaging studies, and laboratory test results. These data may be confounding factors affecting ML results. In addition, because of the complexity and difficulty associated with interpretation of ML algorithms, reproducibility might be hindered. With the development of automatic unsupervised ML models and increased data sharing via electronic medical records, we believe ML and other artificial intelligence technologies will increasingly and rapidly become more feasible and relevant in medicine.

Conclusion

We developed 3 ML algorithms to predict in-hospital mortality among sarcoidosis patients hospitalized for HF. Among the models, RF had the best performance. This study proves the feasibility and applicability of ML techniques in predicting clinical outcomes. However, further studies involving larger datasets with more clinical information would be necessary to improve the algorithm.

Table 2 Evaluation of 3 trained models

Model	Sensitivity	Specificity	AUC (95% CI)
Logistic regression	52.0%	70.0%	0.65 (0.53–0.76)
Random forest	60.0%	66.4%	0.71 (0.59–0.82)
XGBoosting	12.0%	97.2%	0.70 (0.58–0.81)

AUC = area under the receiver operating characteristic curve; CI = confidence interval.

Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosures

The authors have no conflicts to disclose.

Authorship

All authors attest they meet the current ICMJE criteria for authorship.

Patient Consent

This study is based on public database, so no patient consent is needed.

Ethics Statement

Individual patient information in the NIS databases is de-identified, so Institutional Review Board (IRB) approval is not required in studies utilizing this dataset.

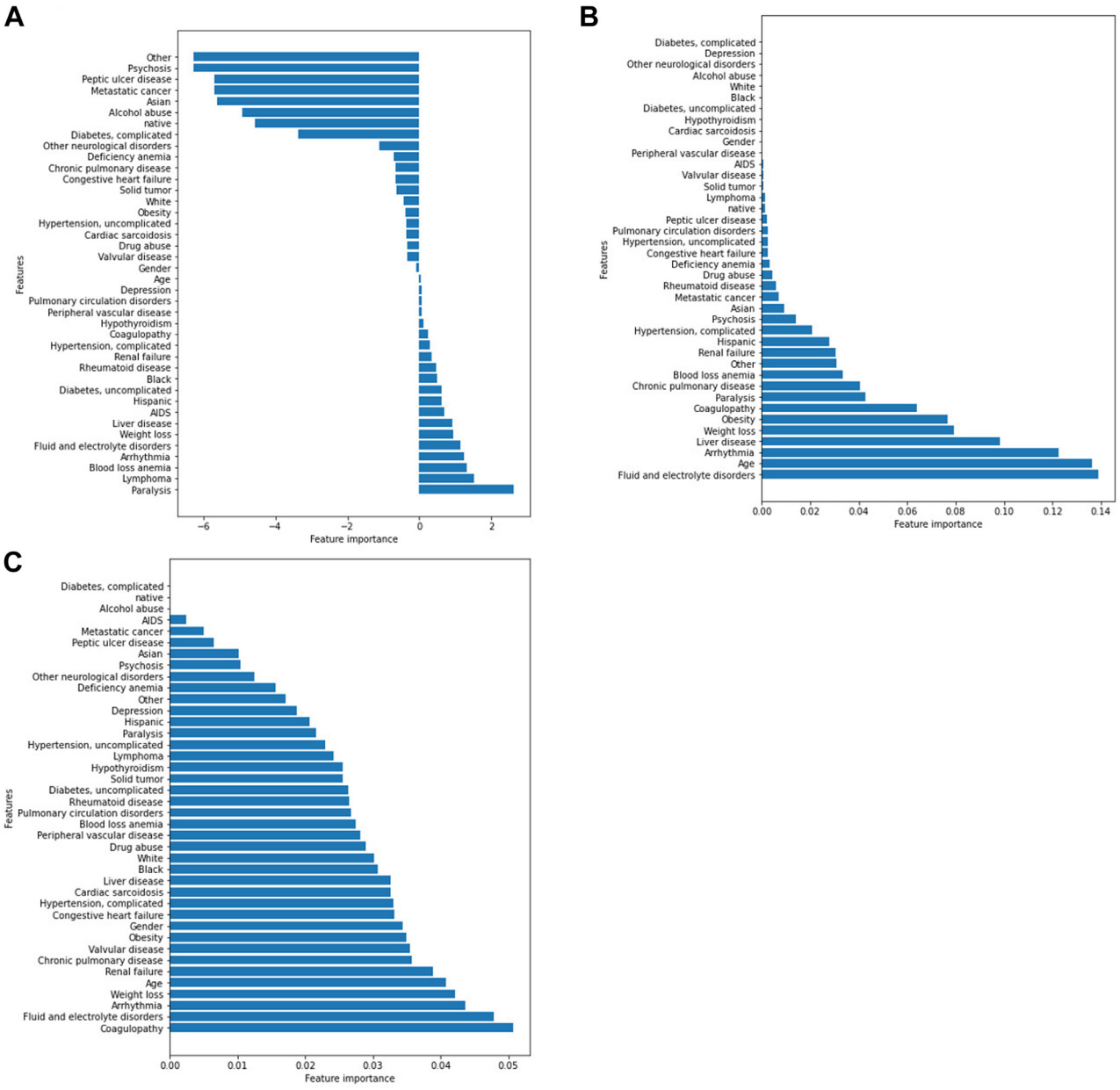

Figure 4 Feature importance of the 3 trained machine learning models. **A:** Logistic regression model. **B:** Random forest model. **C:** XGBoosting model.

Appendix
Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.cvdhj.2022.08.001.

References

1. Patel MR, Cawley PJ, Heitner JF, et al. Detection of myocardial damage in patients with sarcoidosis. Circulation 2009 1977;120:1969.
2. Sekhri V, Sanal S, Delorenzo LJ, Aronow WS, Maguire GP. Cardiac sarcoidosis: a comprehensive review. Arch Med Sci 2011;74:546–554.
3. Houston BA, Mukherjee M. Cardiac sarcoidosis: clinical manifestations, imaging characteristics, and therapeutic approach. Clin Med Insights Cardiol 2014;8(Suppl 1):31–37.
4. Yafasova A, Fosbøl EL, Schou M, et al. Long-term adverse cardiac outcomes in patients with sarcoidosis. J Am Coll Cardiol 2020;76:767–777.
5. Jin CC, Gandrabur L, Kim WY, Pan S, Ash JY. Sarcoid heart disease: review of current knowledge. Cardiol Rev 2021 Jun 14. https://doi.org/10.1097/CRD.0000000000000400.
6. Doughan AR, Williams BR. Cardiac sarcoidosis. Heart 2006;92:282–288.
7. Yazaki Y, Isobe M, Hiroe M, et al. Central Japan Heart Study Group. Prognostic determinants of long-term survival in Japanese patients with cardiac sarcoidosis treated with prednisone. Am J Cardiol 2001;88:1006–1010.
8. Nordenswan HK, Lehtonen J, Ekström K, et al. Manifestations and outcome of cardiac sarcoidosis and idiopathic giant cell myocarditis by 25-year nationwide cohorts. J Am Heart Assoc 2021;10:e019415.
9. Dubrey SW, Bell A, Mittal TK. Sarcoid heart disease. Postgrad Med J 2007;83:618–623.
10. Tellier P, Paycha F, Antony I, et al. Reversibility by dipyridamole of thallium-201 myocardial scan defects in patients with sarcoidosis. Am J Med 1988;85:189–193.
11. Le Guludec D, Menad F, Faraggi M, Weinmann P, Battesti JP, Valeyre D. Myocardial sarcoidosis. Clinical value of technetium-99m sestamibi tomoscintigraphy. Chest 1994;106:1675–1682.

12. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. N Engl J Med 2017;376:2507–2509.
13. Al'Aref SJ, Singh G, van Rosendael AR, et al. Determinants of in-hospital mortality after percutaneous coronary intervention: a machine learning approach. J Am Heart Assoc 2019;8:e011160.
14. Krittanawong C, Virk HUH, Bangalore S, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. Sci Rep 2020;10:16057.
15. Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J 2017;38:500–507.
16. Hernandez-Suarez DF, Kim Y, Villablanca P, et al. Machine learning prediction models for in-hospital mortality after transcatheter aortic valve replacement. JACC Cardiovasc Interv 2019;12:1328–1338.
17. Agency for Healthcare Research and Quality. National (Nationwide) Inpatient Sample (NIS) Database Documentation. Healthcare Cost and Utilization Project (HCUP); 2022.
18. Sharma N, Schwendimann R, Endrich O, Ausserhofer D, Simon M. Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. BMC Health Serv Res 2021;21:13.
19. Moon S, Song HJ, Sharma VD, et al. Classification of Parkinson's disease and essential tremor based on balance and gait characteristics from wearable motion sensors via machine learning techniques: a data-driven approach. J Neuroeng Rehabil 2020;17:125.
20. Maimon O, Rokach L. Data Mining and Knowledge Discovery Handbook. Springer; 2005.
21. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Stat Med 2007;26:5512–5528.
22. Tao C, Huang K, Shi J, Hu Q, Li K, Zhu X. Genomics and prognosis analysis of epithelial-mesenchymal transition in glioma. Front Oncol 2020;10:183.
23. Anderson RP, Jin R, Grunkemeier GL. Understanding logistic regression analysis in clinical reports: an introduction. Ann Thorac Surg 2003;75:753–757.
24. Fawagreh K, Gaber MM, Elyan E. Random forests: from early developments to recent advancements. Syst Sci Control Eng 2014;2:602–609.
25. Liu Y. eXtreme Gradient Boosting (XGBoost): better than random forest or gradient boosting. Yang's Research Blog. 2018. Available at: https://liuyanguu.github.io/post/2018/07/09/extreme-gradient-boosting-xgboost-better-than-random-forest-or-gradient-boosting/. Accessed July 9, 2018.
26. Witten IH, Frank E. Data mining: practical machine learning tools and techniques with Java implementations. ACM Sigmod Record 2002;31:76–77.
27. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics 2013;9:280–299.
28. SekiguchiM, Yazaki Y, Isobe M, Hiroe M. Cardiac sarcoidosis: diagnostic, prognostic, and therapeutic considerations. Cardiovasc Drugs Ther 1996;10:495–510.
29. Ardehali H, Howard DL, Hariri A, et al. A positive endomyocardial biopsy result for sarcoid is associated with poor prognosis in patients with initially unexplained cardiomyopathy. Am Heart J 2005;150:459–463.
30. Iannuzzi MC, Rybicki BA, Teirstein AS. Sarcoidosis. N Engl J Med 2007;357:2153–2165.
31. Kandolin R, Lehtonen J, Airaksinen J, et al. Cardiac sarcoidosis: epidemiology, characteristics, and outcome over 25 years in a nationwide study. Circulation 2015;131:624–632.
32. Cacoub P, Chapelon-Abric C, Resche-Rigon M, Saadoun D, Desbois AC, Biard L. Cardiac sarcoidosis: a long term follow up study. PLoS One 2020;15:e0238391.
33. Patel AR, Rashedi N. Cardiac sarcoidosis: a picture may be worth a thousand words, but do we need more? J Am Heart Assoc 2019;80:e012715.
34. Rojulpote CEA. Predictors of in-hospital mortality in patients with sarcoidosis with cardiomyopathy. Chest 2020;158:A125.
35. Saepudin S, Ball PA, Morrissey H. Hyponatremia during hospitalization and in-hospital mortality in patients hospitalized from heart failure. BMC Cardiovasc Disord 2015;15:88.
36. Brueske B, Sidhu MS, Schulman-Marcus J, Kashani KB, Barsness GW, Jentzer JC. Hyperkalemia is associated with increased mortality among unselected cardiac intensive care unit patients. J Am Heart Assoc 2019;8:e011814.
37. Gologorsky RC, Roy S. Ultrafiltration for management of fluid overload in patients with heart failure. Artif Organs 2020;44:129–139.
38. Zhou Y, Lower EE, Li HP, Costea A, Attari M, Baughman RP. Cardiac sarcoidosis: the impact of age and implanted devices on survival. Chest 2017;151:139–148.
39. Desai R, Kakumani K, Fong HK, et al. The burden of cardiac arrhythmias in sarcoidosis: a population-based inpatient analysis. Ann Transl Med 2018;6:330.
40. Roberts WC, McAllister HA Jr, Ferrans VJ. Sarcoidosis of the heart. A clinicopathologic study of 35 necropsy patients (group 1) and review of 78 previously described necropsy patients (group 11). Am J Med 1977;63:86–108.
41. Sedki M, Fonseca N, Santiago P, et al. Hepatic sarcoidosis: natural history and management implications. Front Med (Lausanne) 2019;6:232.
42. Wand AL, Pavlovic N, Duvall C, et al. Effect of corticosteroids on left ventricular function in patients with cardiac sarcoidosis. Am J Cardiol 2022;177:108–115.
43. Nabeta T, Kitai T, Naruse Y, Taniguchi T, et al. Risk stratification of patients with cardiac sarcoidosis: the ILLUMINATE-CS registry. Eur Heart J 2022 Jul 4:ehac323. https://doi.org/10.1093/eurheartj/ehac323
44. Martusewicz-Boros MM, Boros PW, Wiatr E, Zych J, Piotrowska-Kownacka D, Roszkowski-Śliż K. Prevalence of cardiac sarcoidosis in white population: a case-control study: proposal for a novel risk index based on commonly available tests. Medicine (Baltimore) 2016;95:e4518.