Development and validation of a machine learning-based model for post-sepsis frailty

Hye Ju Yeo [1,2,11], Dasom Noh [3,11], Tae Hwa Kim [1], Jin Ho Jang [1], Young Seok Lee [4], Sunghoon Park [5], Jae Young Moon [6], Kyeongman Jeon [7], Dong Kyu Oh [8], Su Yeon Lee [8], Mi Hyeon Park [8], Chae-Man Lim [8], Woo Hyun Cho [1,2,12] and Sunyoung Kwon [3,9,10,12] on behalf of the Korean Sepsis Alliance investigators

[1]Division of Allergy, Pulmonary and Critical Care Medicine, Department of Internal Medicine, Transplant Research Center, Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, Yangsan, Republic of Korea. [2]Department of Internal Medicine, School of Medicine, Pusan National University, Busan, Republic of Korea. [3]Department of Information Convergence Engineering, Pusan National University, Yangsan, Korea. [4]Division of Respiratory and Critical Care Medicine, Department of Internal Medicine, Korea University, Medical Center, Guro Hospital, Seoul, Republic of Korea. [5]Department of Pulmonary, Allergy and Critical Care Medicine, Hallym University Sacred Heart Hospital, Anyang, Korea. [6]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chungnam National University Hospital, Chungnam National University College of Medicine, Daejeon, Republic of Korea. [7]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea. [8]Department of Pulmonary and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea. [9]School of Biomedical Convergence Engineering, Pusan National University, Yangsan, Korea. [10]Center for Artificial Intelligence Research, Pusan National University, Busan, Korea. [11]H.J. Yeo and D. Noh contributed equally to this article as first authors. [12]S. Kwon and W.H. Cho contributed equally to this article as lead authors and supervised the work.

Corresponding authors: Woo Hyun Cho (chowh@pusan.ac.kr)

Shareable abstract (@ERSpublications)
A model that predicts post-sepsis frailty with high performance was developed using only 10 variables routinely assessed in daily practice. It also performed considerably well in an independent external data set composed of severe COVID-19 patients. https://bit.ly/3URIWhk

Cite this article as: Yeo HJ, Noh D, Kim TH, *et al*. Development and validation of a machine learning-based model for post-sepsis frailty. *ERJ Open Res* 2024; 10: 00166-2024 [DOI: 10.1183/23120541.00166-2024].

Received: 21 Feb 2024
Accepted: 8 May 2024

Abstract

Background The development of post-sepsis frailty is a common and significant problem, but it is a challenge to predict.

Methods Data for deep learning were extracted from a national multicentre prospective observational cohort of patients with sepsis in Korea between September 2019 and December 2021. The primary outcome was frailty at survival discharge, defined as a clinical frailty score on the Clinical Frailty Scale ⩾5. We developed a deep learning model for predicting frailty after sepsis by 10 variables routinely collected at the recognition of sepsis. With cross-validation, we trained and tuned six machine learning models, including four conventional and two neural network models. Moreover, we computed the importance of each predictor variable in the model. We measured the performance of these models using a temporal validation data set.

Results A total of 8518 patients were included in the analysis; 5463 (64.1%) were frail, and 3055 (35.9%) were non-frail at discharge. The Extreme Gradient Boosting (XGB) achieved the highest area under the receiver operating characteristic curve (AUC) (0.8175) and accuracy (0.7414). To confirm the generalisation performance of artificial intelligence in predicting frailty at discharge, we conducted external validation with the COVID-19 data set. The XGB still showed a good performance with an AUC of 0.7668. The machine learning model could predict frailty despite the disparity in data distribution.

Conclusion The machine learning-based model developed for predicting frailty after sepsis achieved high performance with limited baseline clinical parameters.

Introduction

Sepsis-related mortality has steadily declined since the Surviving Sepsis Campaign launch in 2002, and mortality has remained relatively stable [1]. However, the incidence of sepsis amounted to 67.8 per million, and its global burden is still significant [2]. Patients who have had sepsis experience various

clinical courses after discharge; 40% are readmitted within 90 days, one-third die the following year, and one-sixth experience a severe, persistent disability [3]. Only 33% of the patients return to independent living within 6 months after discharge, and 43% return to work within a year [4, 5]. Most of these patients are at an exceptionally high risk of developing long-term sequelae. Approximately half of them present with post-sepsis syndrome, including long-term physical and psychological sequelae [6].

The goal of sepsis treatment should extend beyond survival and improve function after discharge to enhance complete recovery. Identifying individuals at the highest risk of developing functional impairment is critical for facilitating clinical decision-making and providing preemptive care during hospitalisation. Previously, researchers have predicted future functional disability in intensive care unit (ICU) patients [7–10]. However, they did not target patients with sepsis of varying severities or spectra.

Frailty, a measure of functional status, is the leading indicator of several adverse health outcomes derived from the age-related rate of physiological decline. With an increase in interest in frailty, numerous studies have highlighted the impact of pre-existing frailty on sepsis outcomes [11, 12]. Frailty is reversible, and active rehabilitation and nutritional interventions are the preventive and treatment options. Considering the ongoing global medical resource shortage, disparities and negative results of early routine rehabilitation on ICU survival, a focused approach is reasonable. In this context, predicting future frailty is one of the leading priorities in sepsis treatment, focusing on vulnerable populations and improving the efficient distribution of medical resources. Therefore, we aimed to develop and validate a machine learning-based model to predict frailty after sepsis.

Materials and methods

Study population

The Korea Sepsis Alliance (KSA) conducted a nationwide, multicentre, prospective observational cohort study of patients with sepsis between 1 September 2019 and 31 December 2021. The KSA consists of 20 tertiary- and university-affiliated hospitals in South Korea. All patients who visited the emergency departments of participating hospitals and were admitted to general wards or ICU during the study period were screened for eligibility. Patients who developed new sepsis while hospitalised in a general ward were also included in the screening. Adult patients (age $\geqslant$19 years) diagnosed with sepsis, according to the Sepsis-3 definitions, were included and followed up until their hospital discharge or death [13]. Sepsis was diagnosed when infection was presumed or confirmed, and the Sequential Organ Failure Assessment (SOFA) score was $\geqslant$2. Septic shock was diagnosed as patients requiring vasopressors to maintain a mean arterial pressure $\geqslant$65 mmHg despite fluid resuscitation and serum lactic acid level $\geqslant$2 mmol·L^{-1}. The study coordinators at each participating centre prospectively collected data using an electronic case report form (https://sepsis.crf.kr/). They collected the following information: demographics, including age and sex; comorbidities and disease severity (SOFA score, haemodynamics and laboratory variables at baseline); infection source and type (community- and hospital-acquired); multidrug-resistant pathogens in patients with positive cultures; treatment data, such as the adequacy of empirical antibiotic therapy; ICU admissions, resource use; and outcome data, including ICU, 28-day and hospital mortality.

Data collection and extraction for a machine learning-based model

The data for model development and validation were obtained from the KSA Registry. Of the 11 981 patients, 3463 were excluded; 3420 died in the hospital, and 43 were excluded because of missing data (figure 1). We assessed the baseline frailty status before sepsis and status at discharge using the Clinical Frailty Scale (CFS) by medical staff (supplementary table 1) [14]. If the patient had difficulty answering, the guardian confirmed the patient's condition. CFS was measured and collected in 100% of patients at both time points, excluding patients who died. CFS is scored from 1 to 9, with 9 being the maximum score, meaning terminally ill. All patients were classified as "frail" (CFS score 5–9) or "non-frail" (CFS score 1–4) at discharge. The researchers who experimented using machine learning were blinded to the baseline CFS because they developed this model only with CFS at discharge without a baseline CFS. The Institutional Review Board of the Pusan National University Yangsan Hospital waived the requirement for written informed consent and approved this study (approval number 05–2019–092; approval date: 11 July 2019, study title: Korean Sepsis Alliance registry: A multicentre observational cohort study). This study was followed in accordance with the ethical standards of the responsible committee on the Institutional Review Board of the Pusan National University Yangsan Hospital and with the Helsinki Declaration of 1975. Also, this study followed the Guidance for Authors from editors of respiratory, sleep and critical care journals and the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis guidelines [15].

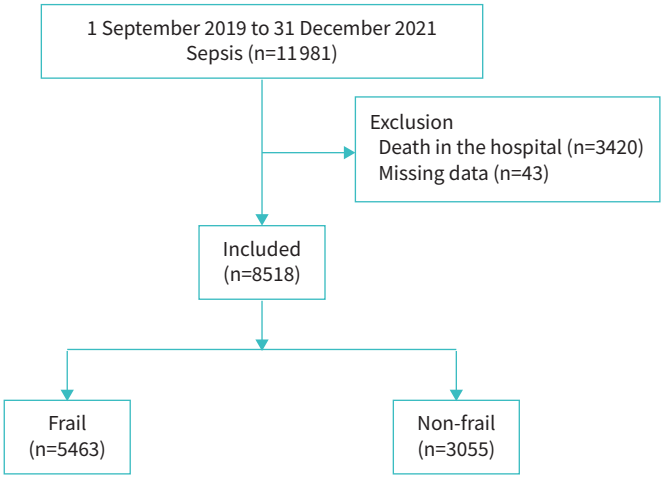

FIGURE 1 Patient inclusion. Overall, 8518 patients were included in this study. At discharge, 5463 (64.1%) and 3055 (35.9%) patients were frail and non-frail, respectively. The patients were divided into frail and non-frail groups according to their Clinical Frailty Scale score at discharge.

Predictors

We selected the predictors based on regression analysis and excluded those with several missing values to obtain sufficient data for model training. We used 10 predictors and data labels for the prediction of post-sepsis frailty (table 1). These 10 predictors consisted of three demographics (age, body mass index (BMI) and initial vital sign heart rate), three initial laboratory findings (Na, Cl and blood urea nitrogen (BUN)), two comorbidities, sources of infection and types of infection according to where the sepsis was acquired (community-acquired, nursing home-acquired, nursing hospital-acquired or hospital-acquired infection).

Machine learning methods, model interpretation and statistical analysis

In this study, we conducted a binary classification to predict frailty at discharge using a few clinical variables at admission. We used seven machine learning models, including six conventional machine learning models, namely, Logistic Regression (LR) [16], Support Vector Machine (SVM) [17], Random Forest (RF) [18], Gradient Boosting Machine (GBM) [19], Extreme Gradient Boosting (XGB) (R package version 0.4–2, 2015) and Balanced Random Forest (BRF) [20], and one neural network model, namely, a Multilayer Perceptron (MLP). XGB and BRF utilised xgboost package [21] and the imbalanced-learn package [22], respectively. Other conventional methods were based on the scikit-learn package [23]. We performed hyperparameter tuning for RF, GBM, XGB and BRF. In RF, the number of trees in the forest and the maximum depth of the tree were set to 100 and 8, respectively. The GBM had a maximum depth limitation of 2, and a subsample ratio of 0.5 was utilised. A total of 50 boosting stages were executed. We tuned an XGB model with a subsample ratio of 0.7, a maximum depth of a tree set to 3 and a minimum sum of instance weight needed in a child set to 5. The model was trained with a learning rate of 0.1. The BRF, consisting of 100 trees, limited the maximum depth of each tree to 7 and required a minimum number of samples of 7 to split an internal node. We constructed a neural network model, the MLP, using the Keras library (Keras, GitHub; https://github.com/fchollet/keras). The MLP consists of two dense layers, each comprising 64 neurons and utilising the “relu” activation function. The output layer employs the sigmoid function.

We conducted a five-fold cross-validation for the experiments. The data set was randomly partitioned into five folds of approximately equal size. The evaluation process is repeated five times, with each fold used as validation and the remaining four folds as training. The results summarised in table 2 represent the mean values of the cross-validation. Additionally, the model was validated through an independent external data set. The external data set retrospectively collected clinical information of severe COVID-19 patients aged 19 years or older with respiratory failure at 22 tertiary hospitals or university hospitals from January 2020 to August 2021 in Korea.

SHapley Additive exPlanations (SHAP) [24] was used to interpret the output of the machine learning model. It provides an explanation by calculating the contribution of each feature to the prediction using

TABLE 1 Patients' characteristics and predictors

Variables	Survivor[#]		Death[§]	p-value	Used features for model
	Non-frail at discharge[¶]	Frail at discharge[+]			
Age years	64.7±14.0	74.4±12.4[f]	72.2±13.2[##,++]		*
Male	1787 (58.5)	3002 (55.0)	2091 (61.1)	<0.001	
BMI kg·m^{-2}	22.8±4.1	21.5±3.9[f]	21.6±4.2[##]		*
Charlson Comorbidity Index	4.7±2.6	5.9±2.4[f]	6.2±2.6[##,++]		
Baseline CFS	3.2±1.4	6.2±1.7[f]	5.7±2.0[##,++]		
Baseline frail	309 (10.1)	4567 (83.6)	2380 (69.6)	<0.001	
Sequential Organ Failure Assessment	5.4±2.7	6.0±2.7[f]	7.7±3.4[##,++]		
Septic shock	489 (16.0)	761 (13.9)	906 (26.5)	<0.001	
Lactate level mmol·L^{-1}	3.1±2.5	3.3±2.8	5.2±4.2[##,++]		
IVS HR rates·min^{-1}	105.9±24.0	104.6±25.4	108.5±26.1[##,++]		*
Na mmol·L^{-1}	134.8±5.9	136.6±8.3[f]	136.0±8.8[##,§§]		*
Cl mmol·L^{-1}	100.9±6.9	102.2±8.8[f]	101.6±9.3[¶¶,ff]		*
BUN mg·dL^{-1}	31.2±23.3	37.3±28.6[f]	44.3±30.1[##,++]		*
Comorbidities					
Diabetes	984 (32.2)	2008 (36.8)	1163 (34.0)	<0.001	
Chronic lung disease	391 (12.8)	772 (14.1)	554 (16.2)	<0.001	
Chronic liver disease	367 (12.0)	376 (6.9)	355 (10.4)	<0.001	
Chronic kidney disease	370 (12.1)	672 (12.3)	481 (14.1)	0.025	
Solid malignancy	1122 (36.7)	1707 (31.2)	1461 (42.7)	<0.001	
Haematological malignancy	192 (6.3)	234 (4.3)	317 (9.3)	<0.001	
Connective tissue disease	106 (3.5)	123 (2.3)	91 (2.7)	0.004	
Cardiovascular disease	621 (20.3)	1305 (23.9)	818 (23.9)	<0.001	*
Chronic neurological disease	269 (8.8)	2008 (36.8)	735 (21.5)	<0.001	*
Sources of infection					*
Pulmonary	795 (26.0)	2663 (48.8)	1769 (51.7)	<0.001	
Abdominal	1278 (41.8)	1210 (22.2)	872 (25.5)	<0.001	
Urinary	579 (19.0)	1371 (25.1)	432 (12.6)	<0.001	
Skin/soft tissue	133 (4.4)	172 (3.2)	88 (2.6)	<0.001	
Catheter-related	32 (1.1)	35 (0.6)	35 (1.0)	0.066	
Neurological	20 (0.7)	43 (0.8)	16 (0.5)	0.195	
Systemic infections without a distinct primary site of infection	312 (10.2)	441 (8.1)	437 (12.8)	<0.001	
Types of infection					*
Community-acquired	2202 (72.1)	2798 (51.2)	1809 (52.9)	<0.001	
Nursing home-acquired	13 (0.4)	575 (10.5)	231 (6.8)		
Nursing hospital-acquired	52 (1.7)	1000 (18.3)	430 (12.6)		
Hospital-acquired	788 (25.8)	1090 (20.0)	950 (27.8)		
Final CFS	6.9±1.0	3.0±0.9[f]			
ICU admission	1128 (36.9)	2039 (37.3)	1702 (49.8)	<0.001	

Data are presented as mean±SD or n (%). BMI: body mass index; CFS: Clinical Frailty Scale; IVS HR: initial vital sign heart rate; BUN: blood urea nitrogen. [#]: n=8518; [¶]: n=3055; [+]: n=5463; [§]: n=3420; [f]: non-frail at discharge *versus* frail at discharge, p<0.001; [##]: non-frail at discharge *versus* death, p <0.001; [¶¶]: non-frail at discharge *versus* death, p <0.01; [++]: frail at discharge *versus* death, p<0.001; [§§]: frail at discharge *versus* death, p<0.01; [ff]: frail at discharge *versus* death, p<0.05.

Shapley Values based on the game theory. TreeExplainer from the Python SHAP package was used to interpret the XGB model. We employed the SciPy package [25] for statistical analysis of patient characteristics. Normality and equal variance tests were first performed before performing each t-test. A chi-square test was performed on categorical variables.

Results

Patient characteristics

The final data set comprised 8518 survival-discharged patients, including 5463 and 3055 patients in the frail and non-frail groups, respectively (table 1). Of those, 3167 were hospitalised in the ICU. There was no difference in rates of ICU admission between the two groups. The mean age of the frail group was

TABLE 2 Performance of seven models for predicting frailty

Models	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score
LR	0.8110	0.7383	0.8288	0.5776	0.7779	0.6530	0.8023
SVM	0.8086	0.7373	0.8250	0.5818	0.7788	0.6496	0.8010
RF	0.8137	0.7371	0.8222	0.5865	0.7802	0.6482	0.8004
GBM	0.8102	0.7364	0.8255	0.5782	0.7776	0.6488	0.8006
XGB	**0.8175**	0.7414	0.8172	0.6070	0.7878	0.6496	0.8020
BRF	0.8136	0.7150	0.6557	0.8207	0.8674	0.5710	0.7468
MLP	0.8139	0.7400	0.8173	0.6027	0.7869	0.6488	0.8011

AUC: area under the receiver operating characteristic curve; PPV: positive predictive value; NPV: negative predictive value; LR: Logistic Regression; SVM: Support Vector Machine; RF: Random Forest; GBM: Gradient Boosting Machine; XGB: Extreme Gradient Boosting; BRF: Balanced Random Forest; MLP: Multilayer Perceptron.

74.4 years, which was significantly higher than that of the non-frail group (64.7 years, p<0.001). The proportion of males in the non-frail group was higher than in the frail group (58.5% *versus* 55%, p=0.002). The SOFA score (5.4 *versus* 6.0, p<0.001) and lactate level (3.1 *versus* 3.3, p=0.005) in the frail group were higher, and the septic shock rate (16% *versus* 13.9%, p=0.009) was lower than in the non-frail group. Regarding comorbidities, the frail group exhibited a higher prevalence of cardiovascular diseases (20.3% *versus* 23.9%, p<0.001) and chronic neurological disorders (8.8% *versus* 36.8%, p<0.001) than the non-frail group. The sites and types of infection differed between the groups. Pulmonary infections were more common in the frail group (26.0% *versus* 48.8%, p<0.001), whereas abdominal infections were more common in the non-frail group (41.8% *versus* 22.2%, p<0.001). In addition, a higher proportion of patients was transferred from nursing homes and hospitals in the frail group (nursing home-acquired 0.4% *versus* 10.5%; hospital-acquired 25.8% *versus* 20.0%, p<0.001).

Model performance

We evaluated and compared the seven models for predicting future frailty using 10 variables at admission, based on their performance metrics, such as area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value and f1 (table 2). All models exhibited great performance, achieving an AUC exceeding 0.8 using only 10 predictors. In particular, XGB achieved the highest AUC (0. 8175) and accuracy (0.7414) for predicting frailty in the sepsis data set. The BRF model, which is specifically designed with a balanced sampling to address imbalanced data, exhibited remarkable performance in terms of PPV. In addition, performance was evaluated for all patients, including those who died (supplementary table S2), and there was no difference from the results including only survivors.

SHAP feature importance

The feature importance analysed using SHAP is depicted in figure 2. Age was the most important variable in frailty prediction, followed by chronic neurological comorbidity and community-acquired infection (figure 2a). Among the sources of infection, pulmonary infection exerted the most influence on the model's predictions, ranking fourth among all predictors. In figure 2b, it is observed that as the feature value of age increases, the corresponding SHAP value also increases. This implies that the model predicts a higher likelihood of frailty in the future as the person's age increases. Similarly, individuals with chronic neurological comorbidity or types of infection classified as nursing home-acquired are predicted to be more likely to be frail at discharge. On the other hand, the model predicts that individuals with community-acquired infections or a higher BMI are less likely to be frail.

External validation with the COVID-19 data set

We conducted external validation to assess the generalisation performance of our models. The models were trained using the entire sepsis data set and validated using an external COVID-19 data set (supplementary table S3). This external validation was challenging due to the disparity in data distribution between the training and validation data sets. The sepsis data set showed differences in all variables between the frail and non-frail groups (p<0.05), while the external validation data set does not in some variables between the two groups for initial heart rate, Na and Cl, with p-values of 0.8826, 0.5858 and 0.1409, respectively. The distribution of labels differs between the two data sets. The prevalence of frail cases in the external validation data set was only 32.80% compared to 64.13% in the sepsis data set (figure 3b,c). This implies that due to the higher proportion of frail cases in the training data, the model may predict and report a more significant number of frail patients during validation. Lastly, because the external validation data set

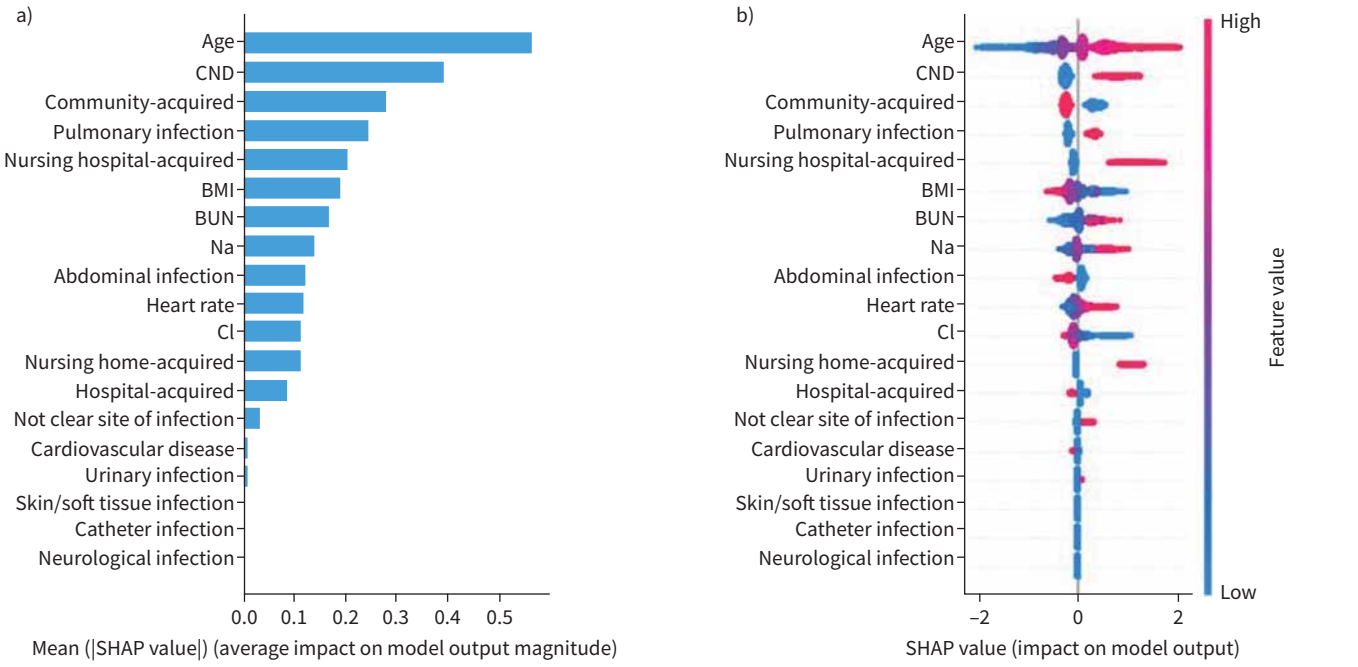

FIGURE 2 Feature importance from SHAP. Priority ranking of risk factors for frailty prediction in the XGB. a) Age has the highest importance in predicting frailty after sepsis. b) As age increases, the possibility of frailty after sepsis increases (red). At a higher BMI, the possibility of frailty after sepsis is lower (blue). CND: chronic neurological disease; BMI: body mass index; BUN: blood urea nitrogen.

consisted only of severe COVID-19 patients, the site of infection was confined to pulmonary infections. Despite these challenging circumstances, the machine learning models showed considerable performance in predicting frailty in a novel data set. The MLP achieved the highest AUC of 0.7668 and an accuracy of 0.7387 (table 3). The XGB model, which demonstrated the highest performance on the sepsis data set, still showed excellent performance (AUC of 0.7619) on the external validation COVID-19 data set (figure 3a).

Discussion

We developed a model that predicts post-sepsis frailty with high performance using only 10 variables routinely assessed in daily practice. Among 8518 patients with sepsis, 64.1% presented with frailty (CFS

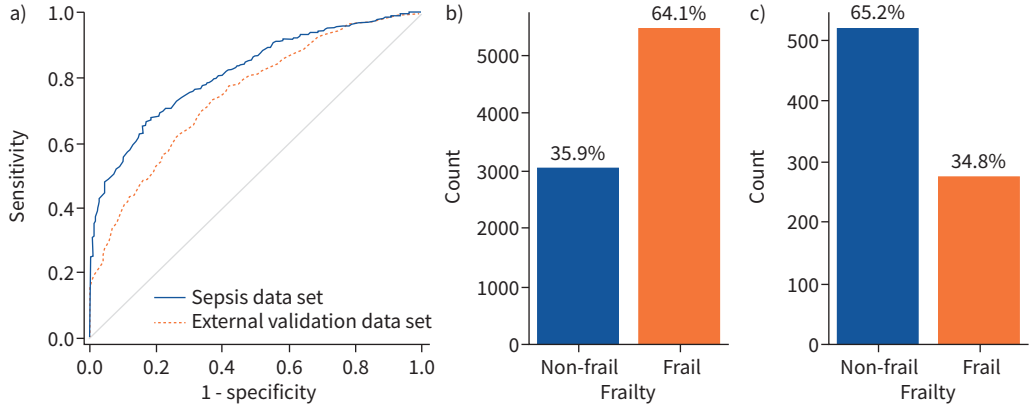

FIGURE 3 Receiver operating characteristic (ROC) curves and the label distribution of the data sets. a) The ROC curves of the Extreme Gradient Boosting (XGB) model in the internal validation (area under the receiver operating characteristic curve (AUC) 0.8175, 95% CI 0.798–0837, p<0.001) data set and external validation data set (AUC 0.7619, 95% CI 0.7275–0.7963, p<0.001); b) the label distribution of the sepsis data set; c) the label distribution of the external validation data set.

TABLE 3 Performance of seven models for predicting frailty in the external validation data set

Models	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score
LR	0.7440	0.7236	0.2599	0.9711	0.8276	0.7109	0.3956
SVM	0.7415	0.7211	0.2419	0.9769	0.8481	0.7071	0.3764
RF	0.7540	0.6344	0.8195	0.5356	0.4850	0.8476	0.6094
GBM	0.7602	0.6533	0.7762	0.5877	0.5012	0.8311	0.6091
XGB	0.7619	0.6771	0.7112	0.6590	0.5267	0.8104	0.6052
BRF	0.7577	0.6960	0.6354	0.7283	0.5552	0.7891	0.5926
MLP	0.7531	0.7073	0.5921	0.7688	0.5775	0.7793	0.5847

AUC: area under the receiver operating characteristic curve; PPV: positive predictive value; NPV: negative predictive value; LR: Logistic Regression; SVM: Support Vector Machine; RF: Random Forest; GBM: Gradient Boosting Machine; XGB: Extreme Gradient Boosting; BRF: Balanced Random Forest; MLP: Multilayer Perceptron.

⩾5) at discharge, and 54.6% got worse CFS scores than before sepsis. The XGB AUC and prediction model accuracy were 0.8175 and 0.7414, respectively. Machine learning predicted whether the baseline CFS before sepsis deteriorated reasonably. This novel study assessed the feasibility of a machine learning method for predicting post-sepsis frailty. This approach will facilitate individualised therapy for sepsis and related physical impairments.

Machine learning approaches in sepsis have been principally used to predict sepsis onset, treatment response, outcomes and phenotype classification [26–31]. The role of machine learning has been expanding to predict long-term outcomes [32]. Physical impairment occurs in numerous patients who have had sepsis, and functional impairment at these recovery points can significantly impact their long-term health and quality of life. The functional outcome at hospital discharge is an essential treatment goal upon considering long-term sequelae and the quality of life as the new goals of sepsis treatment. Clinicians should screen patients at high risk for frailty, evaluate them for common and preventable causes, and modify care delivery to anticipate and diminish these complications [6]. In this cohort, 54.6% got worse CFS scores at discharge, and 10.5% were seriously impacted and finally became frail from non-frail status. Screening and managing groups vulnerable to these dynamic changes in functional status is an essential upcoming issue. This model targets groups vulnerable to functional decline after sepsis without bias. Early management can then focus on these vulnerable and responsive patients. In routine practice, clinicians can easily measure baseline vulnerability through medical history. However, defining the population in which functional decline may occur after sepsis is challenging. This is why we constructed variables without baseline vulnerability to form the prediction model.

Among this population, 53.6% showed persistent frailty both before and after sepsis. (supplementary table S4). In other words, despite efforts, the ABCDE protocol's current practice guidelines regarding ICU rehabilitation may not work for ~53.6% of patients. In this context, clinicians should prioritise the patients who are more vulnerable and may benefit more. One practical difficulty in identifying high-risk groups is the concept of frailty, which is accepted as a phenomenon or comorbid condition and is not an outcome of illness. Although research has brought to our knowledge the comorbidities or demographics associated with frailty, the factors that make patients frail after sepsis are still not known [33]. Therefore, it is only dependent on the physician's experience or the patient's premorbid conditions to screen for high-risk groups of frailty. Thus, it is difficult to standardise preemptive management to prevent frailty. Predictive model development using machine learning can be an alternative solution for optimising decision-making and generating novel insights, leading to more effective and productive actions. Identifying vulnerable patients is the first step to providing optimal treatment to such patients through focused treatment plans, including rehabilitation programmes and dietary prescriptions. Focusing on vulnerable groups can improve the overall outcome of sepsis while ensuring the efficient use of healthcare resources.

Several researchers have developed predictive models for functional impairment in ICU patients [7] [34, 35]. Previous studies targeted heterogeneous patients, and early prediction has limitations. Conversely, Ohbe *et al*. [7] predicted functional disability using 94 clinical indicators within 2 days of admission to the ICU, and the AUC ranged from 0.76 to 0.87. They defined functional impairment as a Barthel Index ⩽60, including severe and total dependency. However, its applicability in real-world ICUs has some drawbacks. First, the Barthel Index is an ordinal scale used to assess performance in the activities of daily living; severely ill patients were not included because the participants were limited to patients who could respond to a self-questionnaire [7]. In addition, they excluded patients with pre-existing frailty by focusing on

recently developed functional disabilities. Second, 94 variables must be collected over 48 h to predict frailty; therefore, the benefits of AI would diminish and be less practical.

We developed a practical model to predict frailty at discharge using a limited number of parameters. Even without baseline CFS before sepsis, it predicted the patients' frailty status and functional decline at discharge well. In particular, it also performed considerably better in an independent external data set composed of severe COVID-19 patients. This study provides novel insights into the baseline predictors of functional outcomes. Previously, in paediatric severe sepsis, some baseline parameters, including the level of organ dysfunction, vasopressor dosing, immune-suppressed state, neurological event, cardiopulmonary resuscitation event or extracorporeal membrane oxygenation use, were associated with post-sepsis frailty [36]. However, adult sepsis was not thoroughly evaluated; it was simply assessed based on demographic factors [37]. In this study, clinically relevant parameters such as initial blood pressure, HCO_3 and lactate, typically indicative of infection severity, did not demonstrate statistical significance concerning post-sepsis frailty. Some of the 10 variables used in our model are clinically relevant, but electrolytes are not fully understood. Currently, the laboratory index for frailty is not yet fully understood either. This will be a new target of future research in frailty from critical illness. Interestingly, the higher the BMI, the lower the probability of frailty in this study; the higher the BUN, the higher the likelihood of frailty. Continuous catabolism leads to severe muscle wasting and associated weakness, resulting in frailty [38]. Thus, patients with high BMI have relatively high reservoirs, consistent with the obesity paradox in sepsis [39]. Moreover, early catabolic markers like BUN may predict functional outcomes after sepsis.

This study had some limitations. First, our results cannot be extended to the impact of predicting frailty at discharge on long-term outcomes because of no information regarding the long-term outcome variables of sepsis. Second, the data sets used for external validation were limited to specific parts of sepsis, such as viral lung infections, and there were also differences in the prevalence of frailty. Despite these limitations, our model showed significant performance in predicting frailty at discharge in external data sets. While our initial target was sepsis, a homogeneous and definitive disease, our model also offers the potential for general applicability to other conditions. This study has widened the scope of machine learning models in critical care, and the developed model is practical and versatile.

Conclusions

Sepsis is a devastating acute disease that results in poor functional outcomes in numerous patients after recovery. We developed and validated a model to predict the risk of frailty after sepsis using 10 common predictors immediately after sepsis awareness. This machine learning-based model achieved high performance with limited baseline clinical parameters. It may facilitate clinical decision-making, initial patient triage, prevention, treatment, and communication with the patients and their families.

Provenance: Submitted article, peer reviewed.

Acknowledgement: We extend our gratitude to the Korean Intensive Care Study Group and the Korean Academy of Tuberculosis and Respiratory Diseases for providing the data set of severe COVID-19 cases. This has enabled us to conduct external validation of our study.

Data availability: The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics statement: The Institutional Review Board of the Pusan National University Yangsan Hospital waived the requirement for written informed consent and approved this study (approval number 05–2019–092; approval date: 11 July 2019).

Author contributions: H.J. Yeo, D. Noh and S. Kwon analysed and interpreted the data. T.H. Kim, Y.S. Lee, S. Park, J.Y. Moon, K. Jeon, D.K. Oh, S.Y. Lee, M.H. Park and C-M. Lee collected the registry. H.J. Yeo and D. Noh were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Korean Sepsis Alliance: Steering committee: Chae-Man Lim (Chair), Sang-Bum Hong, Dong Kyu Oh, Su Yeon Lee, Gee Young Suh, Kyeongman Jeon, Ryoung-Eun Ko, Young-Jae Cho, Yeon Joo Lee, Sung Yoon Lim and Sunghoon Park; participating persons and centres: Jeongwon Heo (Kangwon National University Hospital), Jae-Myeong Lee (Korea University Anam Hospital), Kyung Chan Kim (Daegu Catholic University Hospital), Yeon Joo Lee (Seoul National University Bundang Hospital), Youjin Chang (Inje University Sanggye Paik Hospital), Kyeongman Jeon (Samsung Medical Center), Sang-Min Lee (Seoul National University Hospital), Chae-Man Lim and Suk-Kyung Hong (Asan Medical Center), Woo Hyun Cho (Pusan National University Yangsan Hospital), Sang Hyun Kwak (Chonnam

National University Hospital), Heung Bum Lee (Jeonbuk National University Hospital), Jong-Joon Ahn (Ulsan University Hospital), Gil Myeong Seong (Jeju National University Hospital), Song-I Lee (Chungnam National University Hospital), Sunghoon Park (Hallym University Sacred Heart Hospital), Tai Sun Park (Hanyang University Guri Hospital), Su Hwan Lee (Severance Hospital), Eun Young Choi (Yeungnam University Medical Center), Jae Young Moon (Chungnam National University Sejong Hospital), and Hyung Koo Kang (Inje University Ilsan Paik Hospital).

Conflict of interest: The authors have nothing to disclose.

Support statement: This work was supported by a Research Program funded by the Korea Disease Control and Prevention Agency (fund code 2019E280500, 2021-10-026) and the Korean Sepsis Alliance affiliated with the Korean Society of Critical Care Medicine. This study was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (HI21C1074). This work was supported by Institute of Information and Communications Technology Planning and Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00254177) grant funded by the Korea government (MSIT). Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Prest J, Sathananthan M, Jeganathan N. Current trends in sepsis-related mortality in the United States. *Crit Care Med* 2021; 49: 1276–1284.
2 Rudd KE, Johnson SC, Agesa KM, *et al.* Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet* 2020; 395: 200–211.
3 Prescott HC, Langa KM, Iwashyna TJ. Readmission diagnoses after hospitalization for severe sepsis and other acute medical conditions. *JAMA* 2015; 313: 1055–1057.
4 Yende S, Austin S, Rhodes A, *et al.* Long-term quality of life among survivors of severe sepsis: analyses of two international trials. *Crit Care Med* 2016; 44: 1461–1467.
5 Winters BD, Eberlein M, Leung J, *et al.* Long-term mortality and quality of life in sepsis: a systematic review. *Crit Care Med* 2010; 38: 1276–1283.
6 Prescott HC, Angus DC. Enhancing recovery from sepsis: a review. *JAMA* 2018; 319: 62–75.
7 Ohbe H, Goto T, Nakamura K, *et al.* Development and validation of early prediction models for new-onset functional impairment at hospital discharge of ICU admission. *Intensive Care Med* 2022; 48: 679–689.
8 Schandl A, Bottai M, Holdar U, *et al.* Early prediction of new-onset physical disability after intensive care unit stay: a preliminary instrument. *Crit Care* 2014; 18: 455.
9 Witteveen E, Wieske L, Sommers J, *et al.* Early prediction of Intensive Care Unit-acquired weakness: a multicenter external validation study. *J Intensive Care Med* 2020; 35: 595–605.
10 Wubben N, van den Boogaard M, Ramjith J, *et al.* Development of a practically usable prediction model for quality of life of ICU survivors: a sub-analysis of the MONITOR-IC prospective cohort study. *J Crit Care* 2021; 65: 76–83.
11 Lee HY, Lee J, Jung YS, *et al.* Preexisting clinical frailty is associated with worse clinical outcomes in patients with sepsis. *Crit Care Med* 2022; 50: 780–790.
12 Vanhorebeek I, Latronico N, Van den Berghe G. ICU-acquired weakness. *Intensive Care Med* 2020; 46: 637–653.
13 Singer M, Deutschman CS, Seymour CW, *et al.* The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016; 315: 801–810.
14 Rockwood K, Song X, Macknight C, *et al.* A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005; 173: 489–495.
15 Leisman DE, Harhay MO, Lederer DJ, *et al.* Development and reporting of prediction models: guidance for authors from editors of respiratory, sleep, and critical care journals. *Crit Care Med* 2020; 48: 623–633.
16 Lai TL, Robbins H, Wei CZ. Strong consistency of least squares estimates in multiple regression. *Proc Natl Acad Sci USA* 1978; 75: 3034–3036.
17 Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; 20: 273–297.
18 Breiman L. Random forests. *Mach Learn* 2001; 45: 5–32.
19 Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001; 29: 1189–1232.
20 Chen C, Liaw A, Breiman L, *et al.* Using random forest to learn imbalanced data. University of California, Berkeley, 2004; 110: 1–12.
21 Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International conference on knowledge discovery and data mining. 2016. p. 785–794.
22 Lemaitre G, Nogueira F, Aridas CK. Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. *J Mach Learn Res* 2017; 18: 1–5.
23 Pedregosa F, Varoquaux G, Gramfort A, *et al.* Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011; 12: 2825–2830.

24 Lundberg SM, Lee SI. **Abs./1705.07874** A unified approach to interpreting model predictions. *Arxiv* 2017; preprint [https://doi.org/10.48550/arXiv.1705.07874].
25 Virtanen P, Gommers R, Oliphant TE, *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020; 17: 261–272.
26 Fleuren LM, Klausch TLT, Zwager CL, *et al.* Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med* 2020; 46: 383–400.
27 He Z, Du L, Zhang P, *et al.* Early sepsis prediction using ensemble learning with deep features and artificial features extracted from clinical electronic health records. *Crit Care Med* 2020; 48: e1337–e1342.
28 Schinkel M, van der Poll T, Wiersinga WJ. Artificial intelligence for early sepsis detection: a word of caution. *Am J Respir Crit Care Med* 2023; 207: 853–854.
29 Hu W, Yang M, Chen H. Database-based machine learning in sepsis deserves attention. *Intensive Care Med* 2023; 49: 262–263.
30 Knaus WA, Marks RD. New phenotypes for sepsis: the promise and problem of applying machine learning and artificial intelligence in clinical research. *JAMA* 2019; 321: 1981–1982.
31 Kalimouttou A, Lerner I, Cheurfa C, *et al.* Machine-learning-derived sepsis bundle of care. *Intensive Care Med* 2023; 49: 26–36.
32 Amrollahi F, Shashikumar SP, Meier A, *et al.* Inclusion of social determinants of health improves sepsis readmission prediction models. *J Am Med Inform Assoc* 2022; 29: 1263–1270.
33 Wang X, Hu J, Wu D. Risk factors for frailty in older adults. *Medicine (Baltimore)* 2022; 101: e30169.
34 Haines KJ, Hibbert E, McPeake J, *et al.* Prediction models for physical, cognitive, and mental health impairments after critical illness: a systematic review and critical appraisal. *Crit Care Med* 2020; 48: 1871–1880.
35 Higgins AM, Neto AS, Bailey M, *et al.* Predictors of death and new disability after critical illness: a multicentre prospective cohort study. *Intensive Care Med* 2021; 47: 772–781.
36 Ravikumar N, Sankar J, Das RR. Functional outcomes in survivors of pediatric sepsis: a scoping review and discussion of implications for low- and middle-income countries. *Front Pediatr* 2022; 10: 762179.
37 Iwashyna TJ, Ely EW, Smith DM, *et al.* Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA* 2010; 304: 1787–1794.
38 Gao H, Wang J, Zou X, *et al.* High blood urea nitrogen to creatinine ratio is associated with increased risk of sarcopenia in patients with chronic obstructive pulmonary disease. *Exp Gerontol* 2022; 169: 111960.
39 Yeo HJ, Kim TH, Jang JH, *et al.* Obesity paradox and functional outcomes in sepsis: a multicenter prospective study. *Crit Care Med* 2023; 51: 742–752.