ASTR

ORIGINAL ARTICLE
pISSN 2288-6575 • eISSN 2288-6796
https://doi.org/10.4174/astr.2023.105.4.237
Annals of Surgical Treatment and Research

Explainable machine learning using perioperative serial laboratory results to predict postoperative mortality in patients with peritonitis-induced sepsis

Seung Hee Lim[1], Min Jeong Kim[1], Won Hyuk Choi[1], Jin Cheol Cheong[1], Jong Wan Kim[2], Kyung Joo Lee[3], Jun Ho Park[1,3]

[1]Department of Surgery, Kangdong Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Korea
[2]Department of Surgery, Dongtan Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Korea
[3]Department of Medical Informatics & Statistics, Kangdong Sacred Heart Hospital, Seoul, Korea

Purpose: Sepsis is one of the most common causes of death after surgery. Several conventional scoring systems have been developed to predict the outcome of sepsis; however, their predictive power is insufficient. The present study applies explainable machine-learning algorithms to improve the accuracy of predicting postoperative mortality in patients with sepsis caused by peritonitis.
Methods: We performed a retrospective analysis of data from demographic, clinical, and laboratory analyses, including the delta neutrophil index (DNI), WBC and neutrophil counts, and CRP level. Laboratory data were measured before surgery, 12–36 hours after surgery, and 60–84 hours after surgery. The primary study output was the probability of mortality. The areas under the receiver operating characteristic curves (AUCs) of several machine-learning algorithms using the Sequential Organ Failure Assessment (SOFA) and Simplified Acute Physiology Score (SAPS) 3 models were compared. 'SHapley Additive exPlanations' values were used to indicate the direction of the relationship between a variable and mortality.
Results: The CatBoost model yielded the highest AUC (0.933) for mortality compared to SAPS3 and SOFA (0.860 and 0.867, respectively). Increased DNI on day 3, septic shock, use of norepinephrine therapy, and increased international normalized ratio on day 3 had the greatest impact on the model's prediction of mortality.
Conclusion: Machine-learning algorithms increase the accuracy of predicting postoperative mortality in patients with sepsis caused by peritonitis.
[Ann Surg Treat Res 2023;105(4):237-244]

Key Words: Delta neutrophil index, Machine learning, Peritonitis, Postoperative mortality, Sepsis

INTRODUCTION

Sepsis is one of the most common causes of death in postoperative patients. Despite recent developments in antibiotic treatment and general critical care practices, rates of mortality due to severe sepsis and septic shock are still increasing worldwide [1-3]. In particular, surgical site infections complicate recovery following surgery and increase hospital stay and medical care costs. Some studies have reported that the rate of death from sepsis is between 20% and 30%, and early recognition and risk stratification are needed to improve outcomes for patients with sepsis [4,5].

Received April 7, 2023, **Revised** August 7, 2023, **Accepted** August 28, 2023

Corresponding Author: Jun Ho Park
Department of Surgery, Kangdong Sacred Heart Hospital, Hallym University College of Medicine, 150 Seongan-ro, Gangdong-gu, Seoul 05355, Korea
Tel: +82-2-2224-2222
E-mail: zooono@kdh.or.kr
ORCID: https://orcid.org/0000-0003-3411-8974

Many investigators have tried to identify reliable biomarkers for diagnosis and management of sepsis. CRP and procalcitonin levels are the most frequently used diagnostic and prognostic markers, but their ability is limited [6]. Despite numerous investigations, no single biological marker has been shown to reliably identify patients at risk of developing severe sepsis or septic shock [7,8]. In addition to biomarkers, several clinical scores, including the Acute Physiology and Chronic Health Evaluation (APACHE) score, the Sequential Organ Failure Assessment (SOFA) score, and the Simplified Acute Physiology Score (SAPS), reflect disease severity and predict sepsis outcomes.

Machine learning has been used in various clinical fields in efforts ranging from diagnosis to prediction [9-11]. Because of the success of machine learning in other clinical applications, this study sought to explore whether machine-learning algorithms can also predict mortality in patients with sepsis caused by peritonitis. In addition, explainable artificial intelligence (AI) has recently been introduced to explain the specific decisions made by AI models [12-15]. Explainable AI can help assess the interpretability and explainability of risk predictors for mortality in patients with sepsis caused by peritonitis. This study compares the performance of several machine-learning models with those of conventional SAPS3 and SOFA scores.

METHODS

This study was approved by the Institutional Review Board of the Kangdong Sacred Heart Hospital (No. 2022-03-009). It was performed in accordance with the Declaration of Helsinki and written informed consent was waived due to its retrospective nature.

Patients

We retrospectively reviewed the medical records of patients who underwent surgery for septic peritonitis at the Kangdong Sacred Heart Hospital from July 2011 to December 2021, a 600-bed teaching hospital in Seoul (Fig. 1).

Sepsis is defined as a case in which microbial etiology was proved or suspected in addition to systemic inflammatory response syndrome (SIRS). SIRS is defined by the satisfaction of any 2 of the following components: (a) body temperature of >38 °C or <36 °C; (b) leukocytosis (>10,000/μL), leukopenia (<4,000/μL) or >10% band formation; (c) heart rate of >90 beats/min; and (d) respiratory rate of >24 breaths/min. Septic shock refers to a case in which the arterial blood pressure is continuously <90 mmHg, the mean arterial pressure is <60 mmHg, or a decline in systolic pressure >40 mmHg from baseline occurs without other causes such as hypotension or inadequate fluid resuscitation [16]. Patients aged <18 years; who were pregnant; who had hematologic abnormalities; who had received granulocyte colony-stimulating factors, glucocorticoids, or other immunosuppressants; and who died within 72 hours after surgery were excluded.

Data collection

Clinical characteristics of interest were age, sex, diagnosis, operation time, transfusion, blood culture results, presence of septic shock, and hospital mortality. Postoperative mortality was limited to cases of death within 30 days of surgery due to sepsis only. Patients who required reoperation, mechanical ventilation, or renal replacement therapy were documented as well. The SAPS [17] and SOFA scores [18] were calculated to measure the severity of the patient's condition. The American Society of Anesthesiologists (ASA) physical status classification [19] was referenced from the anesthesia record chart.

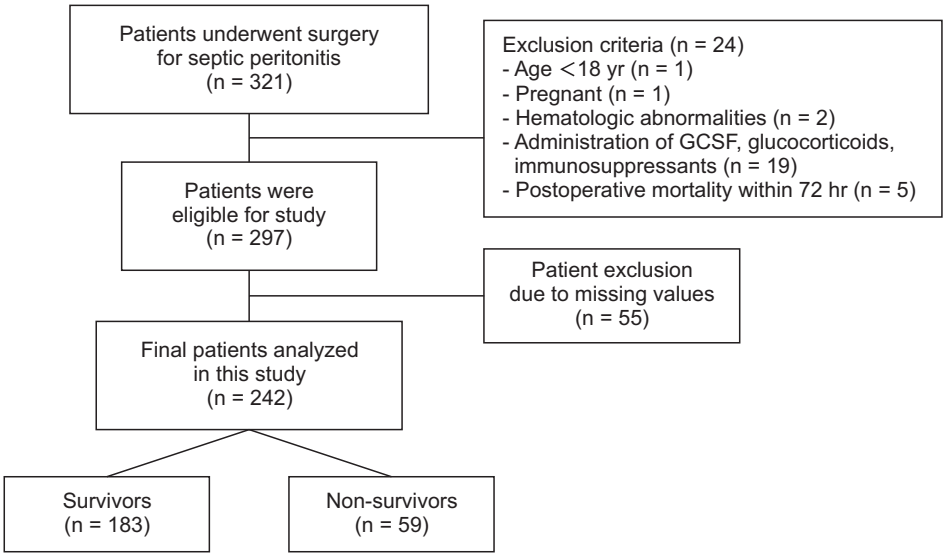

Fig. 1. Patient flowchart. GCSF, granulocyte colony-stimulating factor.

Blood sample measurement

The delta neutrophil index (DNI), WBC and neutrophil counts, and CRP level were measured <12 hours prior to surgery (day 0), 12–36 hours after surgery (day 1), and 60–84 hours after surgery (day 3). DNI measurement is part of the routine complete blood count tests performed at our institution. DNI calculations were performed by an automatic cell analyzer (ADVIA 2120 Hematology System; Siemens Healthcare Diagnostic). DNI values were calculated by subtracting the fraction of polymorphonuclear neutrophils from the number of myeloperoxidase-reactive cells.

Statistical analysis

For statistical analyses, IBM SPSS Statistics ver. 28.0 (IBM Corp.) was used. The categorical variables are shown as absolute and relative frequencies, and continuous variables are shown as mean ± standard deviation values. The chi-square tests were used to compare categorical values, and the independent t-test was used to compare continuous variables between groups. A significant difference was defined as $P < 0.05$.

Table 1. Clinical characteristics of survivors and non-survivors

Variable	Total	Survivors	Non-survivors	P-value
No. of patients	242 (100)	183 (75.6)	59 (24.4)	
Age (yr)	69 (57–79)	67 (59–78)	68 (60–77)	0.484
Male sex	149 (61.6)	112 (61.2)	37 (62.7)	0.849
Body mass index (kg/m^2)	22.7 (19.1–23.9)	22.9 (20.3–24.2)	20.8 (19.2–24.0)	0.538
Septic shock	82 (33.9)	34 (18.6)	48 (74.6)	<0.001
Norepinephrine therapy	52 (21.5)	13 (7.1)	39 (66.1)	<0.001
Surgical site				0.732
Small bowel	77 (31.8)	59 (30.6)	18 (30.5)	
Colon	71 (29.3)	57 (31.1)	14 (23.7)	
Gastroduodenal tract	33 (13.60)	24 (13.1)	9 (15.3)	
Biliary tract	34 (15.0)	24 (14.2)	10 (16.9)	
Others	18 (14.0)	11 (6.0)	7 (11.9)	
Appendix	9 (3.7)	8 (4.4)	1 (1.7)	
Reoperation	25 (10.3)	13 (7.1)	12 (20.3)	<0.010
Operation time (min)	145.0 (108–190)	148.2 (110–190)	150.5 (100–190)	0.945
Positive blood culture	105 (43.4)	78 (42.1)	27 (45.7)	0.131
Mechanical ventilation	169 (69.8)	114 (62.3)	55 (93.2)	<0.001
Renal replacement therapy	25 (10.3)	14 (7.7)	11 (18.6)	0.019
ASA PS score	2.66 ± 0.81	2.58 ± 0.65	2.85 ± 0.93	0.016
SAPS3	66.0 (54–77)	60.9 (52–70)	83.4 (72–97)	<0.001
SOFA score	6.7 (4–9)	5.5 (3–7)	10.7 (8–13)	<0.001
WBC (×10^9/L)				
Day 1	12.2 (6.8–16.0)	12.6 (7.4–16.7)	11.0 (4.8–14.5)	0.200
Day 2	11.0 (6.3–14.3)	11.4 (6.8–14.3)	9.7 (4.9–14.4)	0.130
Day 3	10.5 (7.1–12.5)	10.0 (7.1–12.0)	11.8 (6.6–14.5)	0.021
Neutrophil (%)				
Day 1	80.9 (75.6–90.6)	81.4 (76.0–90.1)	79.5 (70.6–91.9)	0.830
Day 2	83.2 (79.0–91.0)	84.0 (79.8–91.0)	80.5 (75.0–91.0)	0.070
Day 3	82.2 (77.5–89.0)	81.9 (76.8–88.0)	83.1 (78.3–90.8)	0.390
CRP day 3 (mg/L)				
Day 1	109.8 (24.0–165.2)	107.3 (22.4–164.7)	117.5 (30.4–187.2)	0.520
Day 2	150.3 (67.2–210.5)	152.8 (72.3–210.9)	142.5 (58.5–204.4)	0.510
Day 3	139.5 (73.8–184.0)	135.6 (74.8–177.9)	151.5 (64.4–218.9)	0.023
DNI day 3 (%)				
Day 1	15.4 (2.6–20.0)	12.8 (2.3–16.8)	23.3 (5.4–41.5)	<0.001
Day 2	14.0 (2.2–20.1)	10.5 (1.9–11.8)	24.6 (4.6–41.9)	<0.001
Day 3	8.5 (1.2–5.5)	2.8 (1.0–3.3)	25.9 (8.2–39.8)	<0.001

Values are expressed as the or number of patients (percentage).
Values are presented as number (%), mean (interquartile range), or mean ± standard deviation.
ASA, American Society of Anesthesiologists; PS, physical status; SAPS, Simplified Acute Physiology Score; SOFA, Sequential Organ Failure Assessment; DNI, delta neutrophil index.

Machine learning model generation and comparison

Python 3.9.7 (Jupyter Notebook) was used for data handling and machine-learning analysis. The dataset was randomly divided into training (80%) and testing (20%) datasets. Models included in the PyCaret version 3.0 library, such as logistic regression (LR), extreme gradient boost (XGB), gradient boosting classifier, and CatBoost, were used in this study [20]. We performed 10-fold cross-validation on the training set for all models. Through the PyCaret library, machine-learning algorithms were applied to determine the significant models and to fine tune the model hyperparameters.

After the proper models were fitted, their predictive performances were compared using the testing set based on the accuracy, precision, F1 score, and area under the receiver operating characteristic curve (AUC). To calculate the metrics of SAPS3 and SOFA scores, including the accuracy, precision, F1 score, and AUC, we used Youden J statistic and found the optimal thresholds to be 72 and 8, respectively. Interpretation of the prediction model was performed using the SHapley Additive exPlanations (SHAP), a unified approach adopted to precisely calculate the contribution and influence of each feature on the final predictions. The SHAP values show the contribution of each predictor, either positively or negatively, to the target variable. In addition, each observation in the dataset can be interpreted using the set of SHAP values [21].

RESULTS

Baseline clinical characteristics

Table 1 shows the baseline clinical characteristics of the study participants at enrollment. In all, 242 patients were enrolled in this study, including 183 survivors (75.6%) and 59 non-survivors (24.4%). Septic shock, norepinephrine use, reoperation, renal replacement, mechanical ventilator therapy, higher ASA classifications, and higher SAPS3 and SOFA scores were more common among non-survivors than survivors.

The most common cause of reoperation was wound dehiscence, followed by anastomosis leakage, bile leakage, bleeding, intestinal obstruction, and small bowel ischemia.

Mortality prediction model building and evaluation

The AUC values, accuracy, and F1 scores obtained from the test set are listed in Table 2. The AUC value of the CatBoost classifier was 0.933, which was the highest among all the AUCs of machine-learning models (Fig. 2). LR and XGB achieved the next highest AUC values, both being 0.931. Finally, the AUC values of the SOFA scores and SAPS3 for prediction of intensive care unit (ICU) mortality were 0.860 and 0.867, respectively. Overall, the SOFA scores and SAPS3 achieved lower AUC values, accuracies, and F1 scores than the machine-learning models (Fig. 3).

Table 2. Model performances for mortality prediction in the test set

Variable	AUC	Accuracy	F1 score
CatBoost classifier	0.933	0.888	0.748
Logistic regression	0.931	0.876	0.727
Extreme gradient boost	0.931	0.880	0.739
Gradient boosting classifier	0.919	0.888	0.761
Light gradient boosting machine	0.886	0.872	0.724
Random forest classifier	0.869	0.855	0.643
SAPS3	0.860	0.802	0.652
SOFA scores	0.867	0.773	0.626

AUC, area under the receiver operating characteristic curve; SAPS, Simplified Acute Physiology Score; SOFA, Sequential Organ Failure Assessment.

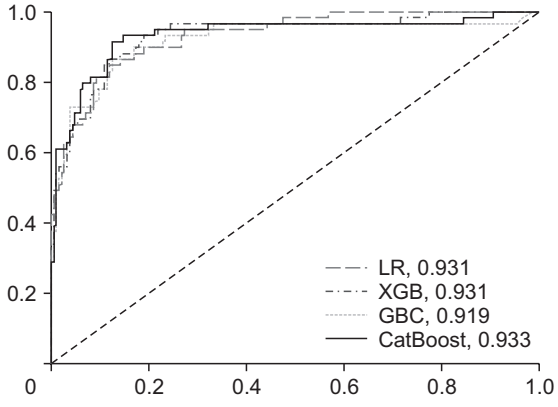

Fig. 2. Receiver operating characteristic curves of the CatBoost classifier, logistic regression (LR), extreme gradient boost (XGB), and gradient boosting classifier (GBC).

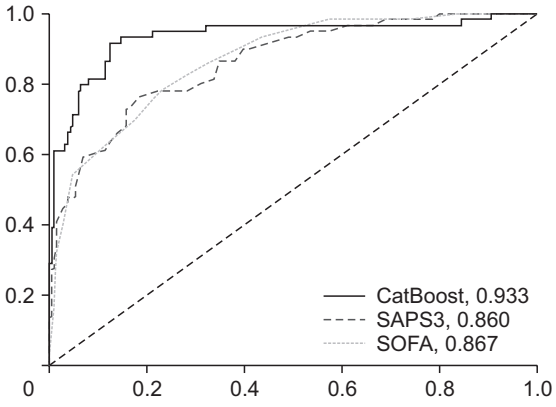

Fig. 3. Receiver operating characteristic curves of the Simplified Acute Physiology Score (SAPS) 3, Sequential Organ Failure Assessment (SOFA) scores, and CatBoost classifier.

Explanation of CatBoost model with SHAP values

The SHAP algorithm was used to determine the significance of each predictor variable to the outcome predicted by the CatBoost model. The SHAP values represent the contribution of a single feature to the prediction of mortality. The variable importance plot lists the most significant variables in descending order (Fig. 4A). The DNI on day 3 had the strongest predictive value for all prediction horizons, followed closely by septic shock and norepinephrine use. For detection of risk factors for mortality, the positive and negative relationships of SHAP values. As presented in Fig. 4B, the SHAP values of all features for all samples were plotted. The horizontal location of features shows whether the effect of a value is associated with a stronger or weaker prediction, and the color shows whether that variable is high (red) or low (blue).

SHAP individual force plots

Fig. 5 shows the individual force plots of a survivor and a non-survivor. The force plot is another way to see the effect of each feature on the prediction for individual patients. In the plot, positive and negative SHAP values are displayed on the left and right sides, respectively. The left (red) side shows the components that increase mortality risk, while the right (blue) side shows those that decrease the mortality risk. The bold numbers are the probabilistic predicted values (f(x)). The f(x) of the survivors was −4.5 and that of the non-survivors was 2.6.

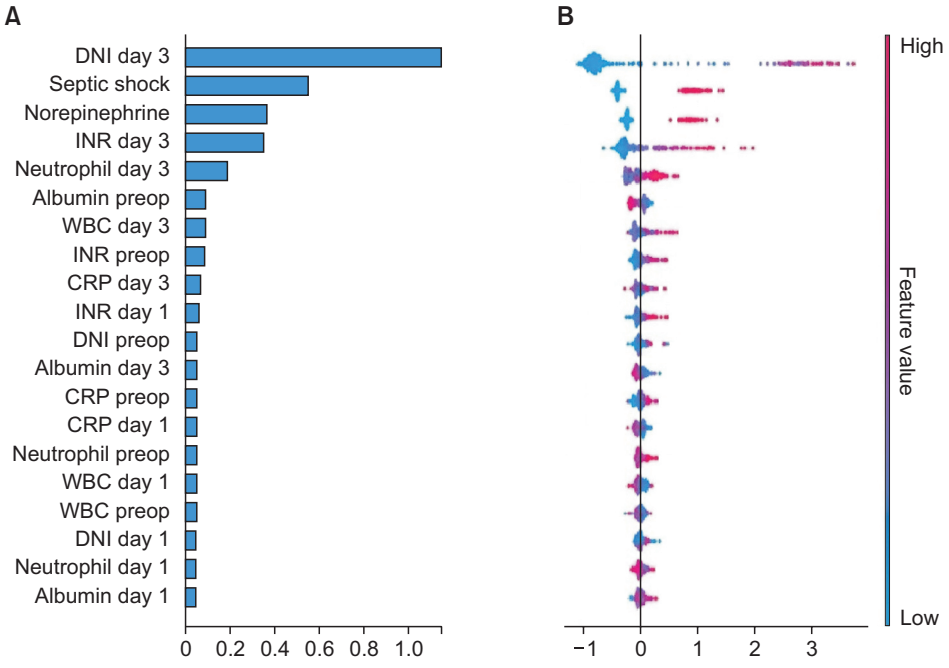

Fig. 4. (A) SHAP feature importance plot. The longer the bar, the larger the impact the feature has on the output. (B) SHAP summary plot. The visualization depicts the influence of a patient's feature value on the prediction through individual dots. SHAP, SHapley Additive exPlanations; DNI, delta neutrophil index; INR, international normalized ratio.

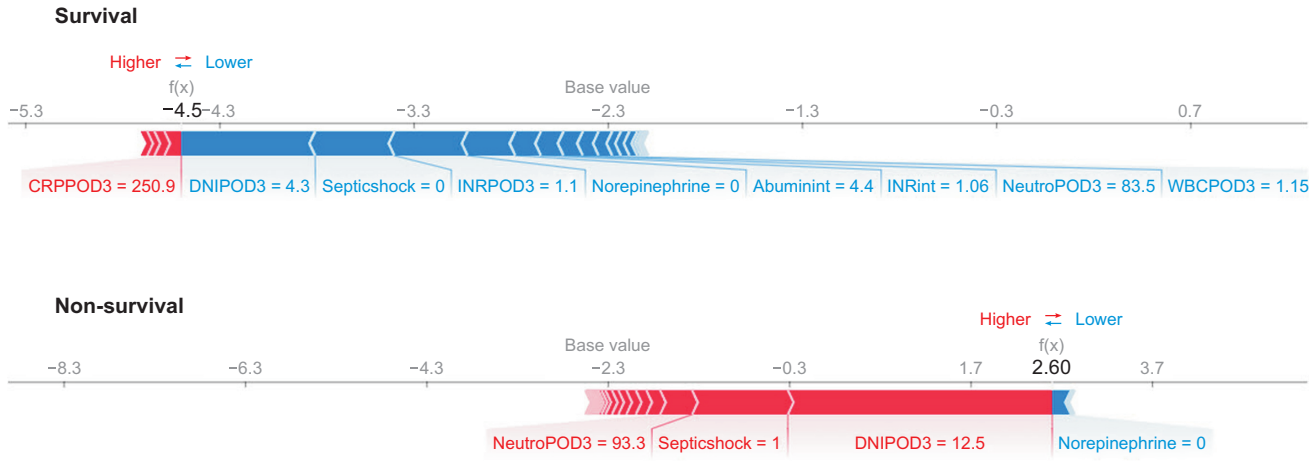

Fig. 5. SHapley Additive exPlanations force plot for 2 selected patients.

The length of the arrow of each feature indicates the degree of contribution. DNI on day 3, septic shock, and norepinephrine use demonstrated a high impact on survivability among both survivors and non-survivors.

DISCUSSION

Sepsis is the leading cause of death in critically ill patients, and its prevalence is increasing globally each year. To improve the treatment outcome of sepsis, early detection and treatment are essential. Biomarkers such as CRP level, procalcitonin concentration, and various cytokines are elevated in sepsis patients and could be useful as prognostic and diagnostic markers during sepsis treatment. Several scoring systems employing clinical and laboratory variables have been developed to predict the outcomes of critically ill patients, including the SAPS3 and SOFA scores. Siwei et al. [22] used the SOFA score and SAPS3 scores to extract mortality predictions after cardiac surgery, with AUC values of 0.809 and 0.850, respectively. Likewise, in a study by Basile-Filho et al. [23] and Pawar et al. [24], the AUC values for the SOFA scores and SAPS3 for postoperative mortality prediction were 0.791 and 0.840, respectively. In several other studies, the classic scoring system showed a similar or lower prediction performance [9,25,26]. For this reason, there is a need to find suitable and specialized prediction modalities for sepsis.

The prediction of mortality using machine learning has become a new field of research. Pirracchio et al. [27] used a composite machine learning-based mortality prediction approach called the Super ICU Learner Algorithm. In this study, the AUC for hospital mortality prediction was 0.88 according to Super Learner compared to 0.71 and 0.78 for SOFA scores and SAPS-II, respectively. With machine learning, Kong et al. [13] predicted in-hospital mortality for sepsis patients in the ICU (AUC, 0.829), and Fransvea et al. [28] predicted mortality for emergency surgery in the elderly (accuracy, 94.9%). Thus, machine-learning algorithms have been applied to predict mortality in sepsis patients, but studies have not focused on their use for patients with sepsis caused by peritonitis.

The present study showed that machine learning can improve the prediction of postoperative mortality in patients with sepsis caused by peritonitis. The AUC of the CatBoost classifier model, the optimal model in this study, was 0.933, which was superior to that of the current clinical scoring system.

The advantage of this study is that explainable machine learning was used to establish a prediction model. Even though machine learning is being used more often in the medical field, black-box issues remain a major concern for physicians when making high-stakes decisions [29]. Lately, explainable AI based on domain knowledge and *post-hoc* analyses has been actively applied to interpret machine-learning models; this could lead to the resolution of black-box issues [12,15]. In this study, the SHAP algorithm was used to interpret the model at the feature level. This leads to a further advantage in that we analyzed several features before and after surgery. The results showed that DNI, INR, CRP, and neutrophil count on day 3 had a greater influence on mortality prediction than did preoperative results.

The WBC and neutrophil percentage can fluctuate in both directions, either increasing or decreasing during sepsis. These bidirectional changes in baseline data can make it challenging to interpret the results accurately. Time series analysis studies on DNI and CRP are quite rare, but in the research conducted by Park et al. [30], they found that the peak time for CRP was 36–50 hours and for DNI was 24 hours. The prompt response of DNI could potentially result in a more accurate depiction of post-surgical sepsis recovery.

These data demonstrate the application of a real dataset and explainable machine-learning approach to establish a physician-understandable postoperative mortality prediction model for patients with sepsis caused by peritonitis.

Our study has several limitations. Since this was a retrospective review of medical records rather than a prospectively designed study, it could only compare WBC count and CRP level, which are measured routinely at the hospital, unlike biological markers such as procalcitonin, lactic acid, and disseminated intravascular coagulation scores. Due to the small sample size in this study, additional studies enrolling larger numbers of patients are required to confirm the best choice of machine-learning model for predicting prognosis after peritonitis surgery and to establish guidelines for sepsis in surgical patients.

Our study found that machine-learning models predict postoperative mortality of patients with sepsis caused by peritonitis better than conventional scoring systems, such as SAPS and SOFA scores. This study indicated that machine-learning algorithms are suitable for clinical use in mortality prediction.

ACKNOWLEDGEMENTS

Fund/Grant Support

This research was supported by a grant from the Kangdong Sacred Heart Hospital Fund (No. 2022-05).

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

ORCID iD

Seung Hee Lim: https://orcid.org/0009-0008-9058-288X
Min Jeong Kim: https://orcid.org/0000-0002-3472-7537
Won Hyuk Choi: https://orcid.org/0000-0002-7194-5330
Jin Cheol Cheong: https://orcid.org/0009-0009-6200-2005

Jong Wan Kim: https://orcid.org/0000-0001-7335-764X
Kyung Joo Lee: https://orcid.org/0000-0002-5965-8064
Jun Ho Park: https://orcid.org/0000-0003-3411-8974

Author Contribution

Conceptualization: JHP
Formal Analysis: JHP, KJL
Methodology: JWK
Project Administration: MJK, WHC, JCC
Writing – Original Draft: JHP, SHL
Writing – Review & Editing: All authors

REFERENCES

1. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med 2001;345:1368-77.
2. Coen D, Cortellaro F, Pasini S, Tombini V, Vaccaro A, Montalbetti L, et al. Towards a less invasive approach to the early goal-directed treatment of septic shock in the ED. Am J Emerg Med 2014;32:563-8.
3. Jo YH, Kim K, Lee JH, Kang C, Kim T, Park HM, et al. Red cell distribution width is a prognostic factor in severe sepsis and septic shock. Am J Emerg Med 2013;31:545-8.
4. Dickinson JD, Kollef MH. Early and adequate antibiotic therapy in the treatment of severe sepsis and septic shock. Curr Infect Dis Rep 2011;13:399-405.
5. Huttunen R, Aittoniemi J. New concepts in the pathogenesis, diagnosis and treatment of bacteremia and sepsis. J Infect 2011;63:407-19.
6. Pierrakos C, Vincent JL. Sepsis biomarkers: a review. Crit Care 2010;14:R15.
7. Kofoed K, Andersen O, Kronborg G, Tvede M, Petersen J, Eugen-Olsen J, et al. Use of plasma C-reactive protein, procalcitonin, neutrophils, macrophage migration inhibitory factor, soluble urokinase-type plasminogen activator receptor, and soluble triggering receptor expressed on myeloid cells-1 in combination to diagnose infections: a prospective study. Crit Care 2007;11:R38.
8. Shapiro NI, Trzeciak S, Hollander JE, Birkhahn R, Otero R, Osborn TM, et al. A prospective, multicenter derivation of a biomarker panel to assess risk of organ dysfunction, shock, and death in emergency department patients with suspected sepsis. Crit Care Med 2009;37:96-104.
9. Kang MW, Kim J, Kim DK, Oh KH, Joo KW, Kim YS, et al. Machine learning algorithm to predict mortality in patients undergoing continuous renal replacement therapy. Crit Care 2020;24:42.
10. Yang F, Wang HZ, Mi H, Lin CD, Cai WW. Using random forest for reliable classification and cost-sensitive learning for medical diagnosis. BMC Bioinformatics 2009;10 Suppl 1(Suppl 1):S22.
11. Obermeyer Z, Emanuel EJ. Predicting the future: big data, machine learning, and clinical medicine. N Engl J Med 2016;375:1216-9.
12. Lauritsen SM, Kristensen M, Olsen MV, Larsen MS, Lauritsen KM, Jørgensen MJ, et al. Explainable artificial intelligence model to predict acute critical illness from electronic health records. Nat Commun 2020;11:3852.
13. Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. BMC Med Inform Decis Mak 2020;20:251.
14. Lin MY, Li CC, Lin PH, Wang JL, Chan MC, Wu CL, et al. Explainable machine learning to predict successful weaning among patients requiring prolonged mechanical ventilation: a retrospective cohort study in central Taiwan. Front Med (Lausanne) 2021;8:663739.
15. Zhang Z, Beck MW, Winkler DA, Huang B, Sibanda W, Goyal H, et al. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. Ann Transl Med 2018;6:216.
16. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Crit Care Med 2003;31:1250-6.
17. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3: from evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med 2005;31:1345-55.
18. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med 1996;22:707-10.
19. Keats AS. The ASA classification of physical status: a recapitulation. Anesthesiology 1978;49:233-6.
20. Moez Ali. PyCaret 3.0: an open-source, low-code machine learning library in Python [Internet]. 2020 [cited Year Mon date]. Available from: https://pycaret.gitbook.io/docs/
21. Scott M. Lundberg S-IL. A unified approach to interpreting model predictions. In: von Luxburg U, Guyon I, Bengio S, Wallach H, Fergus R, editors. NIPS'17: Proceedings of the 31st International Conference on Neural

Information Processing Systems; 2017 Dec 4-9; Long Beach (CA). Curran Associates Inc.; 2017. p. 4768-77.
22. Bi S, Chen S, Li J, Gu J. Machine learning-based prediction of in-hospital mortality for post cardiovascular surgery patients admitting to intensive care unit: a retrospective observational cohort study based on a large multi-center critical care database. Comput Methods Programs Biomed 2022;226:107115.
23. Basile-Filho A, Lago AF, Menegueti MG, Nicolini EA, Rodrigues LA, Nunes RS, et al. The use of APACHE II, SOFA, SAPS 3, C-reactive protein/albumin ratio, and lactate to predict mortality of surgical critically ill patients: a retrospective cohort study. Medicine (Baltimore) 2019;98:e16204.
24. Pawar RD, Shih JA, Balaji L, Grossestreuer AV, Patel PV, Hansen CK, et al. Variation in SOFA (Sequential Organ Failure Assessment) Score Performance in Different Infectious States. J Intensive Care Med 2021;36:1217-22.
25. Kim JW, Park JH, Kim DJ, Choi WH, Cheong JC, Kim JY. The delta neutrophil index is a prognostic factor for postoperative mortality in patients with sepsis caused by peritonitis. PLoS One 2017;12:e0182325.
26. Arts DG, de Keizer NF, Vroom MB, de Jonge E. Reliability and accuracy of Sequential Organ Failure Assessment (SOFA) scoring. Crit Care Med 2005;33:1988-93.
27. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. Lancet Respir Med 2015;3:42-52.
28. Fransvea P, Fransvea G, Liuzzi P, Sganga G, Mannini A, Costa G. Study and validation of an explainable machine learning-based mortality prediction following emergency surgery in the elderly: a prospective observational study. Int J Surg 2022;107:106954.
29. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44-56.
30. Park SJ, Park J, Lee MJ, Seo JS, Ahn JY, Cho JW. Time series analysis of delta neutrophil index as the predictor of sepsis in patients with acute poisoning. Hum Exp Toxicol 2020;39:86-94.