Machine learning model for predicting oliguria in critically ill patients

Yasuo Yamao[1], Takehiko Oami[1], Jun Yamabe[2], Nozomi Takahashi[1], Taka-aki Nakada[1]

Affiliations:

1. Department of Emergency and Critical Care Medicine, Chiba University Graduate School of Medicine, Japan
2. Smart119 Inc., Japan

Corresponding author

Taka-aki Nakada

Department of Emergency and Critical Care Medicine, Chiba University Graduate School of Medicine, 1-8-1 Inohana, Chuo, Chiba, 260-8677, Japan

Phone: +81-43-226-2372, Fax: +81-43-226-2371

Email: taka.nakada@nifty.com

Supplementary materials

Additional File 1:

Table S1. Accuracy and computing time of machine learning classifiers

Table S2. Data source and variables used for machine learning algorithms

Additional File 2:

Figure S1. Flowchart for the enrollment of patients in the study

Additional File 3:

Figure S2. The top 50 important variables in the 1,127-value dataset for predicting oliguria at 6 h

Additional File 4:

Figure S3. SHAP values of the machine learning algorithm for predicting oliguria at 72 h in the ICU

Supplementary Table 1. Accuracy and computing time of machine learning classifiers

Classifier	Computation time (sec)	AUC	F1 score
LightGBM	83	0.984	0.896
CatBoost	454	0.982	0.877
RandomForest	622	0.976	0.842
XGboost	1,480	0.984	0.898

AUC, area under the receiver operating characteristic curve; GBM, gradient boosting machine; CatBoost, Category Boosting; and XGboost, extreme gradient boosting.

Supplementary Table 2. Data source and variables used for machine learning algorithms

Data source	Variables
ICU admission	Age, Sex, SOFA score, APACHE II score
Physiologic measurements	Body mass index, Systolic blood pressure, Mean blood pressure, Diastolic blood pressure, Heart rate, Respiratory rate, oxygen saturation, Central venous pressure, Core temperature, Axillary temperature
Blood tests	AST, ALT, LDH, CPK, Total protein, Albumin, Uremic acid, Urea nitrogen, Creatinine, Total bilirubin, CRP, Interleukin-6, White blood cell, Hemoglobin, Platelet, PT-INR, APTT, FDP, pH, pCO_2, pO_2, HCO_3^-, Base excess, Sodium, Potassium, Chloride, Calcium, Lacatate
Observational record	Urine volume (mL) Urine volume (mL/kg/h) Oliguria
Calculated data	Total amount of input Total amount of output Administration of cardiovascular agents Administration of diuretics Administration of transfusion

SOFA, sequential organ failure assessment; APACHE, acute physiology and chronic health evaluation; AST, aspartate aminotransferase; ALT, alanine transaminase; LDH, lactate dehydrogenase; CPK, Creatinine phosphokinase; CRP, C-reactive protein; PT-INR, prothrombin time-international normalized ratio; APTT, activated partial thromboplastin time; FDP, fibrinogen degradation product.

Figure legends

Supplementary Figure 1. Flowchart for the enrollment of the patients in the study

ICU, intensive care unit.

Supplementary Figure 2. The top 50 important variables in the 1,130-value dataset for predicting oliguria at 6 h

The top 50 important variables in the 1,127-value dataset for predicting oliguria at 6 h are listed in a descending order of importance. Columns with yellow backgrounds overlap with the selected variables in the 50-value dataset.

SOFA, sequential organ failure assessment; MCHC, Mean corpuscular hemoglobin concentration; LDH, lactate dehydrogenase; and FDP, fibrinogen degradation product.

Supplementary Figure 3. SHAP values of the machine learning algorithm for predicting oliguria at 72 h in the intensive care unit

The effect of the features on the output of the model was manifested as the SHAP value. The features are placed in descending order of importance. The correlation between the feature value and SHAP value reveals the beneficial or adverse impact of the predictors. The magnitude of the value is represented by the red (high) or blue (low) graphs.

SHAP, Shapley additive explanations; SOFA, sequential organ failure assessment; FDP, fibrinogen degradation product; LDH, lactate dehydrogenase; and APACHE, acute physiology and chronic health evaluation.