Prediction of fatal adverse prognosis in patients with fever-related diseases based on machine learning: a retrospective study

Chun-Hong Zhao1,2, Hui-Tao Wu3, He-Bin Che3, Ya-Nan Song4, Yu-Zhuo Zhao2, Kai-Yuan Li1, Hong-Ju Xiao2, Yong-Zhi Zhai2, Xin Liu2, Hong-Xi Lu2, Tan-Shi Li1,2

1Medical School of Chinese People’s Liberation Army, Beijing 100853, China; 2Department of Emergency, The First Medical Center of Chinese People’s Liberation Army General Hospital, Beijing 100853, China; 3National Engineering Laboratory for Medical Big Data Application Technology, The First Medical Center to Chinese People’s Liberation Army General Hospital, Beijing 100853, China.

Abstract

Background: Fever is the most common chief complaint of emergency patients. Early identification of patients at an increasing risk of death may avert adverse outcomes. The aim of this study was to establish an early prediction model of fatal adverse prognosis of fever patients by extracting key indicators using big data technology.

Methods: A retrospective study of patients’ data was conducted using the Emergency Rescue Database of Chinese People’s Liberation Army General Hospital. Patients were divided into the fatal adverse prognosis group and the good prognosis group. The commonly used clinical indicators were compared. Recursive feature elimination method was used to determine the optimal number of the included variables. In the training model, logistic regression, random forest, adaBoost, and bagging were selected. We also collected the emergency room data from December 2018 to December 2019 with the same inclusion and exclusion criterion. The performance of the model was evaluated by accuracy, F1-score, precision, sensitivity, and the areas under receiver operator characteristic curves (ROC-AUC).

Results: The accuracy of logistic regression, decision tree, adaBoost and bagging was 0.951, 0.928, 0.924, and 0.924, F1-scores were 0.938, 0.933, 0.930, and 0.930, the precision was 0.943, 0.938, 0.937, and 0.937, ROC-AUC were 0.808, 0.738, 0.736, and 0.885, respectively. ROC-AUC of ten-fold cross-validation in logistic and bagging models were 0.80 and 0.87, respectively. The top six coefficients and odds ratio (OR) values of the variables in the logistic regression were cardiac troponin T (CTnT) (coefficient = 0.346, OR = 1.413), temperature (T) (coefficient = 0.235, OR = 1.265), respiratory rate (RR) (coefficient = –0.206, OR = 0.814), serum kalium (K) (coefficient = 0.137, OR = 1.146), pulse oxygen saturation (SPO2) (coefficient = –0.101, OR = 0.904), and albumin (ALB) (coefficient = –0.043, OR = 0.958). The weights of the top six variables in the bagging model were: CTnT, RR, lactate dehydrogenase, serum amylase, heart rate, and systolic blood pressure.

Conclusions: The main clinical indicators of concern included CTnT, RR, SPO2, T, ALB, and K. The bagging model and logistic regression model had better diagnostic performance comprehensively. Those may be conducive to the early identification of critical patients with fever by physicians.

Keywords: Fever; Infection; Machine learning; Prognosis

Introduction

Fever is the most common chief complaint of emergency patients and is an important pathophysiological process and common symptom of many febrile diseases.1,2 In recent years, major public health events, such as severe acute respiratory syndrome, which are mainly manifested by fever, have attracted worldwide attention.1-3 A fever may occur in sepsis and other infectious diseases, and may also be seen in many non-infectious diseases, such as malignancy, tissue ischemia, cerebrovascular accident, and autoimmune disease.2,6 Sometimes it is difficult to diagnose the cause of fever, such as the fever of unknown origin and the patient’s condition may deteriorate sharply. Early identification of patients at an increasing risk of death may avert adverse outcomes.

Because of the complexity of fever-related illnesses, no biomarker can definitely diagnose sepsis or predict its clinical outcome.1-3 General-purpose illness severity scoring systems such as the Acute Physiology and Chronic Health Evaluation II often contain too many complex...
items or are not specific to people with fever.[8,9] With the continuous development of machine learning technology,[10,11] a machine learning approach has outperformed existing clinical decision rules as well as traditional analytic techniques for predicting in-hospital mortality of emergency department (ED) patients with sepsis.[12] This study, using big data analysis technology, aimed to explore the key factors associated with adverse prognosis of patients with febrile illness, establish an effective model to predict fatal adverse prognosis in patients with febrile disease, and provide technical support for auxiliary clinical diagnosis and treatment decision-making.

Methods

Ethical approval

This was a retrospective study of ED visits. As a retrospective study and data analysis was performed anonymously, this study was exempt from the ethical approval and informed consent from patients.

Study design

This study retrospectively analyzed the clinical data of 28,400 patients admitted to the emergency room from November 2014 to March 2018. Diagnostic criteria: fever was defined as axillary body temperature equal to or greater than 37.3°C. Inclusion criteria was fever (body temperature ≥37.3°C) and age ≥12 years. Exclusion criteria: Patients who died shortly after admission (less than 4 h after admission) and failed to complete laboratory examination. The included patients were divided into the fatal adverse prognosis group and the good prognosis group. Definition: Adverse prognosis group included the patients who experienced cardiopulmonary resuscitation or died during emergency treatment, while good prognosis group included those who did not die during treatment in emergency room and did not undergo cardiopulmonary resuscitation, electric defibrillation, endotracheal intubation, tracheotomy and ventilator assisted respiration.

All data elements for each ED visit were obtained from the Emergency Rescue Database of Chinese People’s Liberation Army General Hospital. After the first measurement of body temperature greater than 37.2°C during the ED visit, sign values and test values were collected. But only the first set of data obtained or generated within 24 h of the ED visit were used as prediction variables. Structured query language queries were written to identify and abstract all demographic information (eg, age and sex) and ED health status (eg, vital signs and laboratory result values). Extraction and screening of variables indicators included vital signs, blood routine, blood biochemistry, coagulation function, and arterial blood gas score analysis, screening of key indicators, and completion of prediction model.

Data analysis

Data cleaning was made on the data obtained according to the inclusion and exclusion criteria. The variables that were excessively missing were removed. The individuals whose data showed obvious errors or missing situation were also removed. A baseline description analysis of the remaining variables was performed, where the normal distribution variables are expressed as mean ± standard deviation; the non-normal distribution variables are expressed using the median (interquartile range). A differential hypothesis test was conducted on the variables. The t-test was used for the numerical variables that did not violate normality test and homogeneity test of variance, while the remaining numerical variables were tested by the Mann-Whitney U test. With α = 0.05 as the test level, the differences of the interested variables were discussed.

In terms of feature selection, recursive feature elimination (RFE) method was used to determine the optimal number of the included variables. The basic idea was that given a specified external learning algorithm, the prediction accuracy of all subsets of variable combinations were calculated through RFE, and number of the subset with the highest prediction accuracy was chosen as the optimal number. Then, the optimal number was used in RFE as the parameter to determine the entry predictor of the final model. In this research, the decision tree was used to select the predictors into the final model. Also, the Pearson correlation coefficient was used to analyze the variable correlation of the factors in model and the results were shown using a heat map. To explore the most important factors, we reduced the optimal number to a smaller one and repeated the above process. The co-existing factors were chosen for further discussion.

We compared the performance of logistic regression, random forest, adaboost and bagging with the same predictors selected above. Accuracy, F1-score, precision, sensitivity and the area under receiver operator characteristic curves (ROC-AUC) were used as criteria for judging model performance. The cohort was split as training and testing set at a ratio of 7:3. The training coefficient of logistic regression was C = 0.01 and the penalty was set to 1.2. The base estimator of bagging was decision tree classifier and the training loss criterion was set to entropy. A total of 1000 trainings were performed. The models with good performance were cross verified by ten folds. For validation, we further collected the emergency room data from December 2018 to December 2019 with the same inclusion and exclusion criterion. The performance was also evaluated by accuracy, F1-score, precision, sensitivity, and ROC-AUC. The descriptive baseline analysis and hypothesis testing were done in IBM SPSS Statistics for Windows (Version 19.0. Armonk, NY: IBM Corp., USA), and the rest of the process was done in Python 3.6.1 (https://www.python.org).

Results

Baseline analysis

The data of 5549 patients with fever were obtained, of which 307 patients had fatal poor outcome event and 5242 had no adverse outcome event. The data contained a total of 70 variables. If certain variable was absent in more than 30% of the patients’ data, it was removed, so the remaining part contained 39 variables. After deleting individuals with missing data, a total of 3474 patients who
### Table 1: Baseline analysis results of 39 variables of 3682 patients with fever.

| Variables                      | Control group ($n = 3474$) | Adverse prognosis group ($n = 208$) | Statistics | $P$   |
|--------------------------------|-----------------------------|-------------------------------------|------------|------|
| Heart rate (beats/min)         | 97 ± (28)                   | 112 ± (36)                          | 7.898*     | <0.001|
| Respiratory rate (breaths/min) | 20.68 (1.00)                | 20.00 (4.75)                        | -2.231     | 0.026|
| SBP (mmHg)                     | 125.00 (31.25)              | 119.00 (42.00)                      | -3.207*    | 0.001|
| DBP (mmHg)                     | 74.00 (149.00)              | 71.00 (26.50)                       | -2.545*    | 0.011|
| SPO2 (%)                       | 98.00 (2.00)                | 98.00 (4.00)                        | -2.748*    | 0.006|
| Temperature (°C)               | 37.70 (0.80)                | 38.00 (1.25)                        | 6.073*     | <0.001|
| CK-MB (U/L)                    | 1.43 (1.94)                 | 3.90 (7.70)                         | 11.561*    | <0.001|
| Direct bilirubin (µmol/L)      | 5.20 (6.30)                 | 6.15 (8.58)                         | 2.288*     | 0.022|
| Total bilirubin (µmol/L)       | 25.34 ± (47.07)             | 27.82 ± (49.95)                     | -0.676*    | 0.499|
| LDH (U/L)                      | 221.70 (161.15)             | 434.25 (452.20)                     | 12.679*    | <0.001|
| AMY (U/L)                      | 49.65 (46.63)               | 93.65 (156.63)                      | 8.228*     | <0.001|
| LIP (U/L)                      | 90.20 (103.63)              | 107.10 (197.43)                     | 1.602*     | 0.109|
| Serum sodium (mmol/L)          | 135.00 (6.50)               | 137.00 (9.42)                       | 3.850*     | <0.001|
| Chloride (mmol/L)              | 102.20 (7.90)               | 101.55 (11.23)                      | 0.000*     | 1.000|
| CTnT (ng/mL)                   | 0.02 (0.03)                 | 0.10 (0.33)                         | 16.121*    | <0.001|
| Serum K (mmol/L)               | 3.90 (0.67)                 | 4.14 (0.94)                         | 5.578*     | <0.001|
| Total protein (g/L)            | 63.20 (13.20)               | 58.65 (14.83)                       | -5.087*    | <0.001|
| Albumin (g/L)                  | 32.80 (8.90)                | 29.90 (8.90)                        | -5.935*    | <0.001|
| ALP (U/L)                      | 70.20 (48.03)               | 80.10 (64.05)                       | 2.464*     | 0.014|
| Uric acid (µmol/L)             | 271.60 (181.63)             | 399.00 (310.00)                     | 9.046*     | <0.001|
| Serum calcium (mmol/L)         | 1.15 (0.36)                 | 1.13 (0.21)                         | -3.479*    | 0.001|
| γ-GT (U/L)                     | 33.35 (62.30)               | 46.15 (84.95)                       | 3.468*     | 0.001|
| Serum creatinine (µmol/L)      | 72.80 (45.23)               | 119.35 (157.98)                     | 9.229*     | <0.001|
| Serum P (mmol/L)               | 0.86 (0.41)                 | 0.96 (0.84)                         | 3.342*     | 0.001|
| BUN (mmol/L)                   | 5.73 (5.08)                 | 10.74 (11.70)                       | 10.543*    | <0.001|
| Creatine kinase (U/L)          | 79.35 (152.85)              | 340.55 (956.04)                     | 11.963*    | <0.001|
| Serum Mg (mmol/L)              | 0.82 (0.14)                 | 0.86 (0.20)                         | 4.208*     | <0.001|
| BNP (pg/mL)                    | 527.40 (1730.63)            | 3694.50 (11226.45)                  | 12.224*    | <0.001|
| ALT (U/L)                      | 18.80 (26.00)               | 39.70 (69.58)                       | 9.014*     | <0.001|
| AST (U/L)                      | 23.70 (30.70)               | 61.65 (181.80)                      | 11.251*    | <0.001|
| Hemoglobin (g/L)               | 111.74 ± (29.17)            | 111.00 ± (31.43)                    | 0.333*     | 0.724|
| RBC (×10^6/L)                  | 3.69 ± (0.95)               | 3.63 ± (1.05)                       | -1.215*    | 0.224|
| PLT (×10^12/L)                 | 178.00 (121.00)             | 144.00 (110.50)                     | -5.124*    | <0.001|
| Hematocrit (L/L)               | 0.33 ± (0.11)               | 0.33 ± (0.13)                       | -0.428*    | 0.669|
| WBC (×10^9/L)                  | 10.07 (7.00)                | 12.70 (8.28)                        | 5.537*     | <0.001|
| Neutrophils (%)                | 0.85 (0.12)                 | 0.87 (0.11)                         | 3.132*     | 0.002|
| Eosinophils (%)                | 0.002 (0.007)               | 0.000 (0.004)                       | -4.499*    | <0.001|
| BASO (%)                       | 0.001 (0.002)               | 0.001 (0.001)                       | -1.831*    | 0.067|
| Monocytes (%)                  | 0.05 (0.04)                 | 0.05 (0.03)                         | -3.040*    | 0.002|

The results were expressed as the median (interquartile range) or mean ± standard deviation. *Z values; †t values. SBP: Systolic blood pressure; DBP: Diastolic blood pressure; SPO2: Pulse oxygen saturation; CK-MB: Creatine kinase myocardial isoenzyme; LDH: Lactate dehydrogenase; AMY: Serum amylnase; LIP: Serum lipase; CTnT: Cardiac troponin T; K: Kalium; ALP: Alkaline phosphatase; γ-GT: γ-glutamyl transpeptidase; P: Phosphorus; BUN: Blood urea nitrogen; Mg: Magnesium; BNP: Brain natriuretic peptide precursor; ALT: Alanine aminotransferase; AST: Aspartate transaminase; RBC: Red blood cell count; PLT: Platelet count; WBC: White blood cell classification count; BASO: Basophilic granulocyte.

The obtained 15 factors through these processes were: heart rate (HR), respiratory rate (RR), systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse oxygen saturation (SPO2), temperature (T), creatine kinase myocardial isoenzyme (CK-MB), total bilirubin (TBIL), lactate dehydrogenase (LDH), serum amylase (AMY), serum lipase (LIP), cardiac troponin T (CTnT), serum calcium (K), total protein (TP), and albumin (ALB). After that, the Pearson correlation test was done, with the details shown in Figure 1. The top six factors correlated with the results were: CTnT (0.260), LDH (0.210), SPO2 (−0.180), CKMB (0.180), HR (0.150), and T (0.150).
Model performance

These factors were then put into logistic regression, decision tree, adaboost, and bagging models were also chosen to select the most suitable algorithm with best performance. In the training model, the accuracy of logistic regression, decision tree, adaboost, and bagging was 0.951, 0.928, 0.924, and 0.924, F1-scores were 0.938, 0.933, 0.930, and 0.930, the precision was 0.943, 0.938, 0.937, and 0.937, the sensitivity was 0.571, 0.524, 0.762, and the ROC-AUC were 0.808, 0.738, 0.736, and 0.885, respectively [Table 2 and Figure 2]. Ten-fold cross-validation was performed on the logistic and bagging models with better comprehensive performance, and their ROC-AUC were 0.80 and 0.87, respectively. In validation part, the decision tree model got the highest accuracy and F1-score, while the bagging model got the highest sensitivity and ROC-AUC.

The coefficients and OR values of the variables in the logistic regression are shown in Table 3. Among them, the top six were: CTnT (coef = 0.346, odds ratio [OR] = 1.413), T (coef = 0.235, OR = 1.265), RR (coef = -0.206, OR = 0.814), K (coef = 0.137, OR = 1.146), SPO2 (coef = -0.101, OR = 0.904), and ALB (coef = -0.043, OR = 0.958). The weights of the variables in the bagging model are shown in Table 3. Among them, the top six were: CTnT (0.221), RR (0.127), LDH (0.088), AMY (0.087), HR (0.058), and SBP (0.047).

To explore the most important factors, we reduced the optimal number of RFE to 11 and repeated the above process. Similarly, the bagging model showed the best performance on sensitivity and ROC-AUC, while the logistic regression showed the best accuracy. The details are shown in Table 2. Ten-fold cross-validation was performed on the logistic and bagging models, and their ROC-AUC were 0.77 and 0.86, respectively. After the number of factors in the model was reduced to 11, the top six coefficients of the logistic model were RR (coef = -0.189, OR = 0.828), SPO2 (coef = -0.140, OR = 0.864),
T (coef = 0.191, OR = 1.210), Na (coef = 0.115, OR = 1.220), CTnT (coef = 0.378, OR = 1.460), and K (coef = 0.200, OR = 1.227). The top six of the bagging model were CTnT (0.251), RR (0.145), P (0.088), ALB (0.711), RBC (0.070), and T (0.070). Also, in validation part, the decision tree model got the highest accuracy and F1-score, while the bagging model got the highest sensitivity and ROC-AUC. The details are shown in Table 2.

**Discussion**

Modern medicine often involves collecting large amounts of physiological data, laboratory results and imaging data into electronic records. The data, however, are complex and multidimensional. It is difficult to find subtle relationships between these data and clinical outcomes using traditional statistical techniques. In this study, the advantage of machine learning is to provide possible innovative solutions for clinical doubts, to find important indicators that may be ignored in treatment for clinicians, and to provide guidance for clinical decision-making to prevent adverse prognosis.

We selected 70 clinical variables associated with “fever” as candidate indicators. It can be seen that in the model with 15 impact factors (HR, RR, SBP, DBP, SPO2, T, CK-MB, TBIL, LDH, AMY, LIP, CTnT, K, TP, and ALB) and in the model with 11 impact factors (RR, SPO2, T, Na, CL, CTnT, K, ALB, Ca, P, and RBC), the co-existing influence

| Variables | Coefficient | OR |
|-----------|-------------|----|
| Cardiac troponin T | 0.346 | 1.413 |
| Respiratory rate | -0.206 | 0.814 |
| LDH | 0.000 | 1.000 |
| Amylase | 0.002 | 1.002 |
| Heart rate | 0.023 | 1.024 |
| Systolic pressure | 0.009 | 1.009 |
| Temperature | 0.235 | 1.265 |
| Albumin | -0.043 | 0.958 |
| Lipase | -0.001 | 1.000 |
| SPO2 | -0.101 | 0.904 |
| Diastolic pressure | -0.012 | 0.988 |
| Total bilirubin | 0.001 | 1.001 |
| Total protein | 0.001 | 0.999 |
| CK-MB | 0.012 | 1.012 |
| Serum kalium | 0.137 | 1.146 |

OR: Odd ratio; LDH: Lactate dehydrogenase; SPO2: Pulse oxygen saturation; CK-MB: Quantitative determination of creatine kinase isoenzyme.
factors were: RR, SPO2, T, CTnT, K, and ALB. These factors could be the main clinical indicators of concern in this experiment, which should be given clinical attention.

Our results suggest that CTnT has the largest weight in the prediction model of adverse prognosis patients with fever, the serum levels of CK, CK-MB, CTnT, and brain natriuretic peptide precursor in patients with fever adverse prognosis group were higher than those in patients with good prognosis group. Further logistic regression analysis showed that serum CK-MB and CTnT levels were independent risk factors for poor prognosis in patients with fever, which should be given clinical attention. Troponin T is one of the clinically recognized markers to reflect myocardial injury. When its level increases, it can reflect the increased severity of myocardial injury, with high specificity.[13-15] Previous studies have suggested that myocardial injury, as reflected by elevated cardiac troponin levels in plasma, is common in patients with community-acquired pneumonia.[16] The main reason for poor prognosis in patients with severe sepsis is that patients often have cardiac dysfunction. Once the cardiovascular system is damaged, the death rate of patients increases greatly.[17]

In this study, ALB was found to be a key factor in predicting the death or poor prognosis in patients with fever. Hoeboer et al.[18] concluded that ALB rather than C-reactive protein may be valuable in predicting and monitoring the severity and course of acute respiratory distress syndrome in critically ill patients with or at risk for the syndrome after new onset fever. Therefore, active correction of hypoproteinemia in patients with fever may improve prognosis.

This study suggested that the vital signs of body temperature, HR, RR, SBP, and DBP had important clinical value in predicting the fatal adverse prognosis in patients with fever, and the sum of their weights was 0.3186. It also showed that SPO2 was the one of the independent protective factors for poor prognosis in patients with fever. When HR, RR, and SBP increased in patients with fever, SPO2 should be detected in time, and should be improved by early non-invasive or invasive mechanical ventilation treatment, so as to reasonably shorten the mechanical ventilation time to reduce the mortality of patients. Single-factor analysis showed that serum sodium of adverse prognosis group was higher than that of good prognosis group, suggesting that for clinical patients with severe fever, timely fluid supplement should be given to correct hyper serum sodium caused by dehydration and other reasons.

In addition, this study suggested that serum K, LDH, AMY, TP, and TBIL were independent factors affecting poor prognosis in patients with fever. In patients with fever, attention should be paid to the changes of the above indicators to save lives.

In the methodological part, when 15 variables were selected, the best results were obtained from the logistic regression model and the bagging model. Logistic regression had the highest accuracy (0.951), while the bagging model had the highest AUC score (0.885). Given the clinical application of the model, the clearer the focus on the indicators, the better it could help doctors make decisions. In addition, if more indicators are involved, it is not always possible to ensure that every factor can be obtained in a short time. Therefore, in the case that the performance of the model does not change much, we believe that the model with fewer factors is more conducive in the clinic. The model may be used to predict, verify and improve future clinical medical practices.

We need to admit that our study has two limitations. First, when we selected cohort, we did not separate people with diversified causes of fever into different groups. Now we cannot calculate the accuracy of different fever causes. This can serve as the direction for our future research. Second, as this was a retrospective study, there were many deficiencies in the data, and the data volume is not very ideal, we hope to have more data in the future.

In conclusion, big data analysis method was adopted to establish a scientific and objective prediction and evaluation model for adverse prognosis of patients with fever in this study. We found the main clinical indicators of concern, and the prediction model established had high diagnostic accuracy and reliability, which may be conducive to the early identification of critical patients with fever by physicians, thus improving the prognosis of patients with fever. The application of big data analysis combined with medical research is helpful to improve the diagnosis and treatment level of febrile critical diseases and the prevention and control of infectious diseases. Research on adverse event prediction model for critical patients with fever quantifies the recognition of critical diseases related to fever and provides a reference model for other similar clinical decision support studies.

Funding
This study was supported by grants from the Big Data R&D Project of the People’s Liberation Army General Hospital (No. 2017MBD-30), and the Science and Technology Innovation Nursery Fund Project of the People’s Liberation Army General Hospital (No. 17KMM50).

Conflicts of interest
None.

References
1. Makoni M, Mukundan D. Fever. Curr Opin Pediatr 2010;22:100–106. doi: 10.1097/MOP.0b013e3283509f95.
2. Walter EJ, Hanna-Jumma S, Carraretto M, Forni L. The pathophysiological basis and consequences of fever. Crit Care 2016;20:200. doi: 10.1186/s13054-016-1375-5.
3. Siu JY. Another nightmare after SARS: knowledge perceptions of and overcoming strategies for H1N1 influenza among chronic renal disease patients in Hong Kong. Qual Health Res 2010;20:893–904. doi: 10.1177/1049732310367301.
4. Goldizen FC. From SARS to avian influenza: the role of international factors in China’s approach to infectious disease control. Ann Glob Health 2016;82:180–188. doi: 10.1016/j.agog.2016.01.024.
5. Tam T. Fifteen years post-SARS: key milestones in Canada’s public health emergency response. Can Commun Dis Rep 2018;44:96–101. doi: 10.14745/ccdr.v44i05a01.

6. Kaul DR, Flanders SA, Beck JM, Saint S. Brief report: incidence, etiology, risk factors, and outcome of hospital-acquired fever: a systematic, evidence-based review. J Gen Intern Med 2006;21:1184–1187. doi: 10.1111/j.1533-1477.2006.00566.x.

7. Hotchkiss RS, Moldawer LL, Opal SM, Reinhart K, Turnbull IR, Vincent JL. Sepsis and septic shock. Nat Rev Dis Primers 2016;2:16045. doi: 10.1038/nrdp.2016.45.

8. Smith GB, Peytherche DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013;84:463–470. doi: 10.1016/j.resuscitation.2012.12.016.

9. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med 2015;7:299ra122. doi: 10.1126/scitranslmed.aab3719.

10. Scott DJ, Lee J, Silva I, Park S, Moody GB, Celi LA, et al. Accessing the public MIMIC-II intensive care relational database for clinical research. BMC Med Inform Decis Mak 2013;13:9. doi: 10.1186/1472-6947-13-9.

11. Sandfort V, Johnson AEW, Kunz LM, Vargas JD, Rosing DR. Prolonged elevated heart rate and 90-day survival in acutely ill patients: data from the MIMIC-III database. J Intensive Care 2019;6:622–629. doi: 10.1177/2051165218736828.

12. Taylor RA, Pare JR, Venkatesh AK, Mowaфи M, Melnick ER, Fleischman W, et al. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. Acad Emerg Med 2016;23:269–278. doi: 10.1111/acem.12876.

13. Shah M, Parmaks M, Maladum O, Patel B, Trispathi B, Agarwal M, et al. Mortality in sepsis: comparison of outcomes between patients with demand ischemia, acute myocardial infarction, and neither demand ischemia nor acute myocardial infarction. Clin Cardiol 2018;41:936–944. doi: 10.1002/clc.22978.

14. Masson S, Caironi P, Fanizza C, Carrer S, Caricato A, Fassini P, et al. Sequential N-terminal pro-B-type natriuretic peptide and high-sensitivity cardiac troponin measurements during albumin replacement in patients with severe sepsis or septic shock. Crit Care Med 2016;44:707–716. doi: 10.1097/CCM.0000000000001473.

15. Sasaki B, Butz T, Prull MW, Liebeton J, Christ M, Trappe HJ. Earliest bedside assessment of hemodynamic parameters and cardiac biomarkers: their role as predictors of adverse outcome in patients with septic shock. Int J Med Sci 2015;12:680–688. doi: 10.7150/ijms.11720.

16. Frencken JF, Van Raai L, Kappen TH, Donker DW, Horn J, van der Poll T, et al. Myocardial injury in critically ill patients with community-acquired pneumonia: a cohort study. Ann Am Thorac Soc 2019;16:606–612. doi: 10.1513/AnnalsATS.201804-286OC.

17. Honda T, He Q, Wang F, Redington AN. Acute and chronic remote ischemic conditioning attenuate septic cardiomyopathy, improve cardiac output, protect systemic organs, and improve mortality in a lipopolysaccharide-induced sepsis model. Basic Res Cardiol 2019;114:15. doi: 10.1007/s00395-019-0724-3.

18. Hoeboer SH, Straaten OV, Groeneveld AJ. Albumin rather than C-reactive protein may be valuable in predicting and monitoring the severity and course of acute respiratory distress syndrome in critically ill patients with or at risk for the syndrome after new onset fever. BMC Pulm Med 2015;15:22. doi: 10.1186/s12890-015-0015-1.

How to cite this article: Zhao CH, Wu HT, Che HB, Song YN, Zhao YZ, Li KY, Xiao HJ, Zhai YZ, Liu X, Lu HX, Li TS. Prediction of fatal adverse prognosis in patients with fever-related diseases based on machine learning: a retrospective study. Chin Med J 2020;133:583–589. doi: 10.1097/CM9.0000000000000675