Original Research Article

Artificial intelligence in critical care: prediction of sepsis in patients in intensive care from first initial laboratory parameters

Samrat Mitra*, H. Sangwan

Department of Medicine, Military Hospital Jalandhar Cantt, Punjab, India

*Correspondence:
Dr. Samrat Mitra,
E-mail: samratmitra79@gmail.com

ABSTRACT

Background: Sepsis is a leading cause of morbidity and mortality in the critical care setting. The analysis of hemostatic parameters at admission have been proven to be a predictive marker for development of sepsis in the ICU. The present study aims to develop a machine learning model which can predict the development of sepsis after 72 hours of ICU admission, from initial assessment of hemostatic parameters.

Methods: A total of 170 ICU admissions over six months (May 2018 - Dec 2018) period were included in the study. Hemostatic parameters including platelet counts, prothrombin time and Sonoclot assay were assayed at time of admission. The patients were followed up for development of sepsis. The data was split in two sets: training (100) and test (70). A machine learning model was developed using the linear discriminant analysis (LDA) model, in the R programming environment. The statistical parameters employed were sensitivity, specificity, positive and negative predictive value.

Results: A comparison of incidence of development of clinical sepsis and predicted sepsis by the model showed 74.19% sensitivity and 84.61% specificity over the testing set. 06 false positives and 08 false negative predictions were encountered.

Conclusions: The model shows potential to be used as a predictive tool for development of sepsis in the critical care ward. Moderate sensitivity and good specificity were achieved by the model, highlighting the role of hematologic assessment at admission in prediction of development of sepsis. However, further studies with larger datasets are required before implementation in clinical practice.

Keywords: Artificial intelligence, Coagulation study, D-dimer, Fibrinogen, Machine learning, Sepsis, Sonocolot

INTRODUCTION

Sepsis has emerged as a major challenge for clinicians, managers and healthcare policymakers, and presents with a high fatality rates in critical care setting. It is important to identify markers for an early diagnosis of sepsis and organ dysfunction. In recent years, there has major revisions to our in our understanding of the factors which lead to development of sepsis. The role of impaired hemostasis in development of sepsis has been well documented. Recent advancements in the study of coagulation have elucidated the important contribution of cells to the hemostatic process. The cell-based model places emphasis on platelets and tissue factor-bearing cells while also taking into account the contribution of membrane surfaces, microparticles, enzyme systems, and endothelial cells. An in-depth review of the cell-based model of coagulation has been published previously.[1] In 1889 Hayem suggested that quantification of the changes that occur in blood viscosity during clotting could be utilized as the basis for a test that monitors coagulation function in patients with sepsis.[2] In addition, elasticity of

a blood clot, which is affected primarily by fibrin and platelets in the sample, also plays a role in the development of sepsis.[3,4]

In the light of these new findings, monitoring of coagulation is important to diagnose potential to develop sepsis.[5] Our current understanding of in vivo coagulation highlights the limitations of standard coagulation tests, such as prothrombin time (PT) and activated partial thromboplastin time (aPTT), which do not incorporate cellular elements or only provide data on isolated components of the coagulation cascade. Routine laboratory-based coagulation tests (e.g., prothrombin time/International Normalized Ratio, activated partial thromboplastin time, fibrinogen) and platelet numbers are being used to assess the patients' current coagulation status. However, the value of these tests has been questioned in the acute perioperative setting because there are delays from blood sampling to obtaining results (45-60 min), coagulation tests are determined in plasma rather than whole blood, no information is available on platelet function (PF) and the assays are performed at a standard temperature of 37°C rather than the patient's temperature.[6]

Point-of-care (POC) coagulation monitoring devices assessing the viscoelastic properties of whole blood, such as Sonoclot analysis may overcome several limitations of routine coagulation tests.[7] Blood is analyzed at the bedside and not necessarily in the central laboratory, allowing faster turnaround times. Sonoclot analysis, including the three parameters clot rate (CR), platelet function (PLTFN), activated clotting time, have been found to be predictive of sepsis in previous studies.[8]

The emergence of Artificial Intelligence has revolutionised the practice of medicine, both in the clinical and laboratory setting. Statistical models have been developed to predict patient outcomes in critical care settings.[9] A large meta analysis revealed the impact of machine learning algorithms in critical care setting. Machine learning algorithms have successfully nullified inter clinician variability in the critical care, as well as development of sepsis and targeted therapy.[10-12]

The aim of the present study is to develop a machine learning algorithm to predict the development of sepsis after 72 hours of ICU admission, based on a select set of predictive tests at admission.

METHODS

Inclusion criteria

A total of 170 admissions in the critical care ward of a tertiary care hospital between May 2018 - Dec 2018 were included in this study. All patients were within the age group 20 years to 60 years of age. All the patients included in the study had a minimum stay of 72 hours in the ICU.

Exclusion criteria

Patients (1) having prior bleeding or coagulation disorders, or those on anticoagulant therapy, were excluded from the study, (2) who have received blood transfusion within 48 hours before ICU admission or requiring blood transfusion at admission, (3) patients with deranged liver function tests at the time of admission, (4) patients with vascular thrombosis at any site during admission, were excluded from the study.

Study population includes the patients belonged to the North Indian populace, in and around the state of Punjab.

Initial assessment

Hematological and biochemical parameters were assessed immediately on admission; the following parameters were selected for this study.

- Total leukocyte count (TLC), measured in a Sysmex XP100 3 part hematology analyser.
- Platelet count (PLT), measured with Beckman Coulter LH750 5 part hematology analyser.
- International normalised ratio (INR), measured with ECL 760 coagulation analyser.
- Serum fibrinogen (FIB), measured with ECL 760 coagulation analyser.
- Serum D-dimer (DD) measured with ECL 760 coagulation analyser.
- Activated partial thromboplastin time (APTT), measured with ECL 760 coagulation analyser.
- Sonocolot parameters – clot rate (CR), platelet function (PLTFN), activated clotting time (ACT), measured with a Sienco single channel SC1 sonoclot coagulation analyser.

After the initial measurement, patients were treated as per critical care protocols in the Intensive Care Unit (ICU). After 72 hours of ICU admission, the patients were assayed for presence of sepsis with clinical and laboratory criteria. The criterion for diagnosis of sepsis was[13]

- Temperature < 96.8 °F or >100.4 °F
- Heart rate >90/min
- Respiratory rate >20/min or PaCO2< 32 mmHg
- WBC <4000/mm³ or > 12,000/mm³, or 10% bands
- Suspicion or evidence of focus of bacterial infection

The data was tabulated and split into two subsets

- A 'training' set of 100 records.
- A 'test' set of 70 records.

The machine learning model was trained with the training dataset of 100 records. First, the variables were plotted against each other to test for redundancy. The 'screeplot'

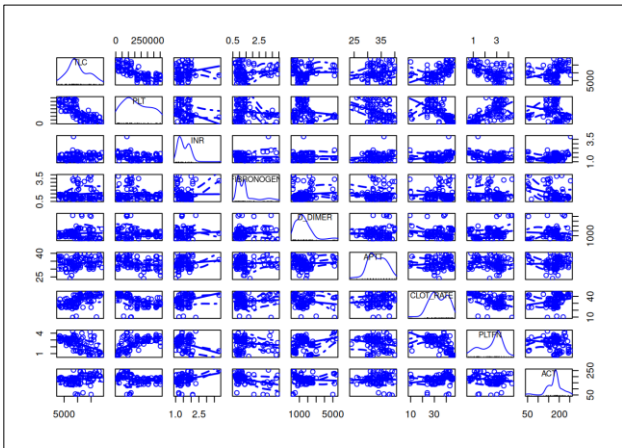method of the R platform was used for this purpose (Figure 1).

Figure 1: Screeplot of predictive variables of the study.

With no redundancy detected, a principal component analysis (PCA) was carried out, which revealed 10 principal components; Figure 2 shows the principal components in descending order of variance. The plot was generated with the PCA function of the R platform.
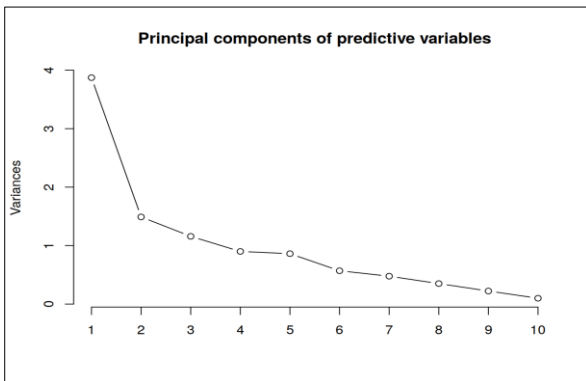

Figure 2: Principal component of predictive variables.

The principal components were then used for a Linear Discriminant Analysis (LDA) which produced a clear demarcation between two groups, sepsis and no sepsis (Figure 3). The LDA score for patients who did not develop sepsis (group 0) was between -4 to 0, whereas the ones who developed sepsis had a score between 0 to +6.

The LDA model was then used to predict the outcomes from the 'test' datatset of 70 records. The model was developed in the R programming language and Rstudio statistical analysis software.[14]
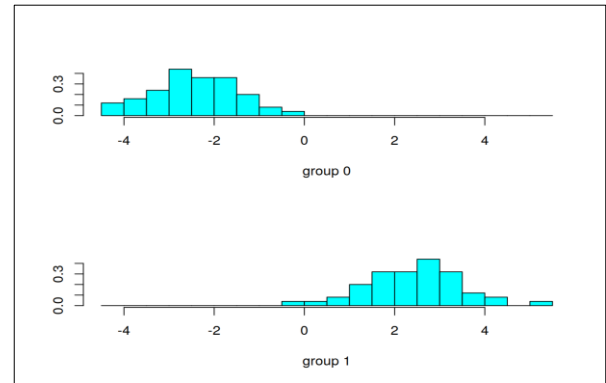

Figure 3: Separation of sepsis (1) and no sepsis (0) by the machine learning model.

RESULTS

The predictions of 70 records from the 'test' dataset were as Table 1. These 70 records were chosen randomly from the master dataset using the random number generator function in R.[14] The machine learning model was run on these 70 records to predict the outcome, i.e. development of sepsis after 72 hours (outcome = 1), or no development of sepsis (outcome = 0). A list of predictions was produced using the predict() function in R. Standard statistical parameters such as sensitivity, specificity, positive predictive value and negative predictive value were calculated from the contingency table (Table 1).

Table 1: Performance of the model on the test set.

		Actual diagnosis		
		Sepsis developed	Sepsis not developed	Total
Predicted by the model	Sepsis predicted	23	6	29
	Sepsis not predicted	8	33	41
	Total	31	39	70
Sensitivity	23/21 = 74.19%	Specificity	33/39 = 84.61%	
Positive predictive value	23/29 = 79.31%	Negative predictive value	33/41 = 80.48%	

In 31 patients who developed actual sepsis after 72 hours of admission, the model could correctly predict the

outcome in 23 cases (74.19% sensitivity). 08 cases were falsely predicted not to develop sepsis (false negatives).

Of 39 patients who did not develop sepsis, 06 were predicted by the model to develop sepsis (false positives). The model showed 74.19% sensitivity and 84.61% specificity. The positive predictive value was a moderate 79.31% and negative predictive value was 80.48%. The slightly higher negative predictive value indicates the hematological and sonoclot testing to have value in screening for development of sepsis at ICU admission.

The results indicate the potential utility of a machine learning model to successfully predict the outcome of patients admitted in critical care ward, from the initial hematological tests as well as sonoclot analysis. Based on the findings of the study, the recommendation for hematological and sonoclot screening of all critical care patients at admission can be reinstated. The study has ruled out majority of false positive and false negative cases by carefully selecting the stud population. All patients with prior bleeding disorders, coagulopathies, or any factor that could affect their coagulation parameters, were excluded from the study. Thus, the bias from prior coagulation abnormalities was ruled out.

DISCUSSION

The present study demonstrates the close correlation between altered hemostatis parameters and the development of sepsis in an ICU setting, as well as the capability of a machine learning model to successfully predict the development of sepsis based on hemostatic parameters.

Previous studies have demonstrated an AUC (area under the curve) of 0.78, an error rate of 0.24, a sensitivity of 0.65 and a specificity of 0.80 in prediction of sepsis.[12] Nemati et al to train their model, they used data from more than 31,000 admissions to the ICUs at two Emory University hospitals. To test their model, they used data from over 52,000 ICU patients from the publicly available Medical Information Mart for Intensive Care-III database. Their model hit its highest prediction performance level when predicting sepsis 4 h prior to its occurrence (AUC 0.85).[15]

Machine learning models have also been used to predict common post-surgical complications.[16] The present study achieves moderate sensitivity and specificity (74.19% and 84.61%, respectively) in identifying patients at risk of developing sepsis. However, the positive and negative predictive value were not satisfactory enough to employ the tool in clinical decision making at this stage.

A number of false positives (06 cases) and false negatives (08 cases) sprang up in the study, which might be attributed to overfitting by the model to the training data. A study with a larger and more varied sample will successfully mitigate this issue.

REFERENCES

1. Smith SA. The cell-based model of coagulation. J Emerg Crit Care. 2009;19:3-10.
2. Hayem G. Du Sang et de Ses Alterations Aanatomique [Blood and its Anatomical Alterations]. Paris, France: G. Masson; 1889;323.
3. Hartert H, Schaeder J. The physical and biological constants of thromboelastography. Biorheol. 1962; 1:31-9.
4. Kaibara M. Dynamic viscoelastic study of the formation of fibrin networks in fibrinogen-thrombin systems and plasma. Biorheol. 1973;10:61-73.
5. Ferraris VA, Ferraris SP, Saha SP, Hessel II EA, Haan CK, Royston BD, et al. Perioperative blood transfusion and blood conservation in cardiac surgery: the Society of Thoracic Surgeons and The Society of Cardiovascular Anesthesiologists clinical practice guideline. Ann Thorac Surg. 2007 May 1;83(5):S27-86.
6. Kozek-Langenecker S. Management of massive operative blood loss. Minerva Anesthesiol 2007;73:401-15.
7. Luddington RJ. Thrombelastography /thromboelastometry. Clini Lab Haematol. 2005 Apr;27(2):81-90.
8. Ding R, Wang Z, Lin Y, Liu B, Zhang Z, Ma X. Comparison of a new criteria for sepsis-induced coagulopathy and International Society on Thrombosis and Haemostasis disseminated intravascular coagulation score in critically ill patients with sepsis 3.0 a retrospective study. Blood Coagul Fibrinolysis. 2018;29(6):551-8
9. Shillan D, Sterne JA, Champneys A, Gibbiso n B. Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. Crit Care. 2019 Dec 1;23(1):284.
10. Lovejoy CA, Buch V, Maruthappu M. Artificial intelligence in the intensive care unit. Crit Care. 2019;23(1):7.
11. Vellido A, Ribas V, Morales C, Sanmartín AR, Rodríguez JC. Machine learning in critical care: state-of-the-art and a sepsis case study. Biomed engineering online. 2018 Nov 1;17(1):135.
12. Pirracchio R, Cohen MJ, Malenica I, Cohen J, Chambaz A, Cannesson M, et al. Big data and targeted machine learning in action to assist medical decision in the ICU. Anaesthes Crit Care Pain Med. 2019 Aug 1;38(4):377-84.
13. Lever A, Mackenzie I. Sepsis: definition, epidemiology, and diagnosis. BMJ. 2007 Oct 25;335(7625):879-83.
14. R: What is R? Available at: https://www.r-project.org/about.html. Accessed 29 June 2019.
15. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG, et al. An interpretable machine

learning model for accurate prediction of sepsis in the ICU. Crit Care Med. 2018;46:547-53.

16. Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, Sündermann SH, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med. 2018 Dec 1;6(12):905-14.