

RESEARCH

Open Access



Development and validation of a machine learning-based model to assess probability of systemic inflammatory response syndrome in patients with severe multiple traumas

Alexander Prokazyuk^{1*} , Aidos Tlemissov² , Marat Zhanaspayev³ , Sabina Aubakirova³  and Arman Mussabekov³ 

Abstract

Background Systemic inflammatory response syndrome (SIRS) is a predictor of serious infectious complications, organ failure, and death in patients with severe polytrauma and is one of the reasons for delaying early total surgical treatment. To determine the risk of SIRS within 24 h after hospitalization, we developed six machine learning models.

Materials and methods Using retrospective data about the patient, the nature of the injury, the results of general and standard biochemical blood tests, and coagulation tests, six models were developed: decision tree, random forest, logistic regression, support vector and gradient boosting classifiers, logistic regressor, and neural network. The effectiveness of the models was assessed through internal and external validation.

Results Among the 439 selected patients with severe polytrauma in 230 (52.4%), SIRS was diagnosed within the first 24 h of hospitalization. The SIRS group was more strongly associated with class II bleeding (39.5% vs. 60.5%; OR 1.81 [95% CI: 1.23–2.65]; $P=0.0023$), long-term vasopressor use (68.4% vs. 31.6%; OR 5.51 [95% CI: 2.37–5.23]; $P<0.0001$), risk of acute coagulopathy (67.8% vs. 32.2%; OR 2.4 [95% CI: 1.55–3.77]; $P<0.0001$), and greater risk of pneumonia (59.5% vs. 40.5%; OR 1.74 [95% CI: 1.19–2.54]; $P=0.0042$), longer ICU length of stay (5 ± 6.3 vs. 2.7 ± 4.3 days; $P<0.0001$) and mortality rate (64.5% vs. 35.5%; OR 10.87 [95% CI: 6.3–19.89]; $P=0.0391$). Of all the models, the random forest classifier showed the best predictive ability in the internal (AUROC 0.89; 95% CI: 0.83–0.96) and external validation (AUROC 0.83; 95% CI: 0.75–0.91) datasets.

Conclusions The developed model made it possible to accurately predict the risk of developing SIRS in the early period after injury, allowing clinical specialists to predict patient management tactics and calculate medication and staffing needs for the patient.

Level of evidence Level 3.

Trial registration The study was retrospectively registered in the ClinicalTrials.gov database of the National Library of Medicine (NCT06323096).

*Correspondence:
Alexander Prokazyuk
prokazyuk.md@yandex.ru

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Keywords SIRS, Machine learning, Polytrauma, Decision making

Introduction

Balancing the timely provision of intensive care with early surgical intervention is difficult in managing patients with severe polytrauma because of its complex pathophysiology [1–4]. Hasty surgical intervention against the background of an acute homeostasis disorder leads to an increased risk of death due to the aggravation of concomitant complications of trauma [5]. In turn, excessive postponement of surgery can lead to late complications such as bone fusion disorders, sepsis, thrombosis, or late death [6, 7]. In both cases, the treatment of such patients leads to excessive economic costs and the excessive use of clinical resources [8–10].

Systemic inflammatory response syndrome (SIRS) is a formidable complication of acute trauma. It is a hyperergic reaction of the immune system to stress factors for the localization and elimination of an endogenous or exogenous source of damage. Despite the initial protective mechanism, the cytokine storm underlying the pathogenesis of SIRS can cause a massive inflammatory cascade, leading to reversible or irreversible dysfunction of internal organs [6, 11–13]. The incidence of SIRS and its frequent complications (sepsis, acute renal injury, and multiple organ dysfunction syndrome) among patients with polytrauma increases with age [14–17]. In addition, SIRS is one of the main predictors of mortality in polytrauma patients. Baek et al. reported that the presence of two clinical signs of the syndrome is significantly associated with mortality [12]. According to a survey of doctors by the European Society of Trauma and Emergency Surgery, among respondents, the fact that a patient with an injury has SIRS is a reason for postponing secondary (definitive) surgery in 67.2% of cases [18].

The prediction of clinical outcomes is a complex task owing to indirect correlations between health indicators and the clinical picture, which are usually left unnoticed by doctors [19]. The use of machine learning (ML) in such cases is beneficial. ML models can easily operate with a large amount of data and identify nonlinear relationships between clinical features [20]. The ease of use of the final product is reliable and increases the clinical effectiveness of young professionals [21].

Our study aimed to create an ML model capable of accurately predicting the risk of developing SIRS 24 h after admission to the emergency department, which is a barrier to early surgical intervention in patients with severe polytrauma. This model potentially allows the identification of patients who may benefit from additional preoperative preparations. In addition, reasoned planning of the management scheme will reduce the

financial costs of treating the case and reduce the burden on the staff.

Methods

Patient selection

Patients were selected through a retrospective-prospective study of electronic medical records (EMRs) of patients who were urgently hospitalized for multiple injuries between January 2018 and January 2024. Study data were collected from the EMRs of two regional medical centers and the “National Scientific Center of Traumatology and Orthopedics” in the Republic of Kazakhstan (the service populations are 610, 1135, and 1400 thousand people, respectively).

We used the Berlin definition of polytrauma [22], with a change in severity criteria, as the basis to identify suitable patients. The original definition requires an Injury Severity Score (ISS) greater than 15 points, which in a minimal form can be represented as one “serious” (AIS post-dot code 3) and two “moderate” (AIS post-dot code 2) injuries. However, given the superiority of the New Injury Severity Score (NISS) over the ISS in terms of improved accuracy and simplified scoring [23], the NISS scale was used as the final evaluation tool. Thus, the study included adult patients (≥ 18 years old) who met both expanded criteria:

1. Severity of injury according to NISS > 16 points.
2. The presence of one or more physiological risk factors and/or primary hospitalization in the ICU. Physiological risk factors are represented by the following indicators: systolic arterial pressure (SAP) ≤ 90 mmHg; Glasgow Coma Scale (GCS) score ≤ 8 points; base excess ≤ -6.0 mmol/L; international normalized ratio (INR) ≥ 1.4 or activated partial thromboplastin time (APTT) ≥ 40 s; and age ≥ 70 years.

The obligatory inclusion criterion was the completeness of the EMRs in terms of laboratory and instrumental studies, and the protocol of therapeutic and surgical treatment. The EMRs were considered acceptable in the presence of up to 10% empty values for the features of interest.

The following patients were excluded: those seeking primary care 24 h after the injury; those who were transferred or required transfer between departments and hospitals for rehabilitation or other stages of therapy; those with prematurely interrupted treatment; those with an injury combined with suffocation, drowning, frostbite, electrical trauma, chemical and/or thermal burns; those

with pathological fractures; pregnant women; and those with dominant severe cranioccephalic (GCS \leq 7 points for more than 5 days) or spinal injury (deep paresis and plegias).

Systemic inflammatory response syndrome

We determined the presence of SIRS in patients if any two of the listed clinical signs [24] manifested continuously for more than 4 h in the first 24 h after admission, and were related to the injury.

- Body temperature $< 36\text{ }^{\circ}\text{C}$ or $> 38\text{ }^{\circ}\text{C}$.
- Heart rate > 90 beats/min.
- Respiratory rate > 20 per minute or $\text{pCO}_2 < 32$ mmHg.
- White blood cell count $> 12 \times 10^9/\text{L}$, $< 4 \times 10^9/\text{L}$, or $> 10\%$ of immature forms.

Data collection

At the time of admission of each patient to the emergency department, health indicators were evaluated and entered into the EMRs by qualified medical personnel. The volume of visible injuries was assessed by the senior doctor on duty in the emergency room, with the involvement of other specialists in conducting additional instrumental and laboratory examination methods. Based on the severity of the injuries and the clinical data obtained, patients were moved along one of the following routes: to the polytrauma department; to the antishock hall with subsequent redetermination of the route; to the ICU; to the operating unit.

Clinical features were collected from anonymized EMRs and divided into two groups.

1. *Baseline variables* – patient data collected upon admission to the emergency department, which included sex; age; list of Abbreviated Injury Scale (AIS update 2008) confirmed codes [25]; NISS; vital signs during hospitalization (pulse, blood pressure, respiratory rate, body temperature); clinical tests: general blood tests (hemoglobin, erythrocytes, hematocrit, platelets, leukocytes); biochemical blood parameters (total protein, total bilirubin, glucose, urea, creatinine); and coagulation tests (INR, APTT, fibrinogen).
2. *Outcome variables* – features characterizing the development of the following complications in a patient: SIRS, acute traumatic coagulopathy (ATC), bleeding over 25% of the circulating blood volume (BV), a new case of pneumonia, the outcome of hospital stay, the length of hospital stay (LOS), and the length of stay in the ICU after admission (ICU LOS).

Data preprocessing

First, the AIS scores were transformed into nine new features corresponding to AIS anatomical areas. The square root of the AIS severity of each confirmed injury was calculated and summed with the values of other existing injuries for this anatomical area. Next, there was a multivariate imputation of the missing values because of all the baseline variables values ($n=8770$); only 33 (0.38%) were missing. We used iterative imputation based on linear regression (all missing values were quantitative), using an ascending order of imputation based on the k-nearest neighbors. Additionally, the values for MCH, MCHC, and MCV were calculated in accordance with generally accepted formulas. Using the reference values from our laboratory, additional binary features reflecting the decrease and excess of laboratory values were generated. Then, low-informative synthetic variables were removed. Thus, the final data table contained 60 features for baseline variables and 7 features for outcome variables (a total of 67 features).

Finally, the database was divided into three sets of data: training, internal, and external validation (Fig. 1). First, 25% of the randomly selected patients were included in the external validation set. The remaining data were divided at a ratio of 3:1 into training and internal validation sets (furtherly defined as a “development set” when mentioned together). An independent value normalization strategy (MinMaxScaler, Python Scikit-Learn library) for quantitative features was applied to each of the three sets. Although the number of cases in the database was relatively small, we did not use any synthetic techniques to increase or decrease the datasets. Only stratification based on the SIRS values was used.

Model development

To predict the risk of developing SIRS within 24 h after admission to the emergency department, six models were developed: decision tree classifier (DTC), random forest classifier (RFC), gradient boosting classifier (GBC), and support vector classifier (SVC), logistic regression (LR) and neural network (black box classification, BBC). When building decision trees (DTC, RFC) and training SVC, parameter autotuning with cross-validation (Scikit-Learn GridSearchCV) was used. For LR and GBC, recursive feature elimination with cross-validation (Scikit-Learn RFECV) was used to select the most significant features. The dimensionality of the neural network was reduced recursively by removing the features that brought the least weight to the overall result of the model until the best result in the test.

Graphs and performance indicators were obtained based on the results of the developed models. All models were tested on an external validation sample with a re-evaluation of the performance indicators. The choice of

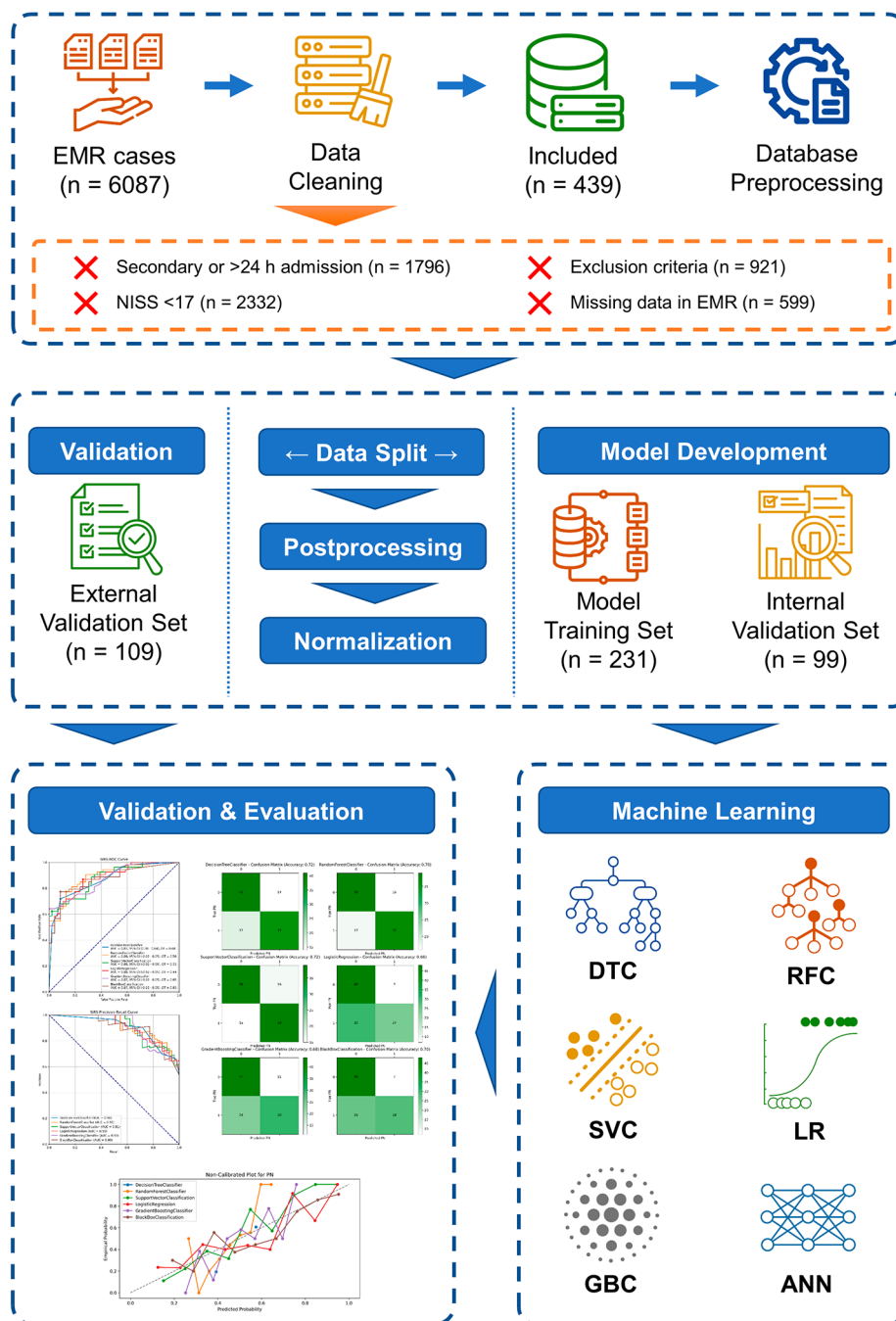


Fig. 1 Schematic representation of the study. Study design and machine learning process. The experimental design included data collection, data pre-processing, data splitting, model development, model validation and performance evaluation. Various modeling techniques have been utilized, including decision tree classification (DTC), random forest classification (RFC), support vector classification (SVC), logistic regression (LR), gradient boosting classification (GBC), and artificial neural network (ANN) methods

the best model was based on a scoring system [26] that included 11 metrics (precision, recall, and F1-score by class; accuracy, overall false-positive rate (FPR) and true-positive rate (TPR), area under receiver operating characteristic curve (AUROC), and area under precision-recall curve (AUPRC)), each of which was rated from 1 to 6,

where the highest score correlated with the best predictive ability. The scoring system ranged from 0 to 66. This system was used twice, during development and external validation. Then, both score tables were summarized in a weight ratio of 0.33 from the development table and 0.67 from the validation table. The model with the highest

final score was considered the most predictably effective. The SHapley Additive exPlanation (SHAP) method was used to analyze the importance of the features in the best model.

Statistical analysis

Continuous data are expressed as numbers (n), percentages (%), mean values (μ), standard deviations (SD), and interquartile ranges (IQRs). Categorical data are expressed as frequencies and percentages. Numeric variables were tested with an independent samples t-test, and categorical variables were tested with Pearson's chi-squared test. All tests were two-sided, and statistical significance was set at $P < 0.05$.

To assess the performance of the models, we used specificity, sensitivity, precision, recall, F1-score, accuracy, overall FPR, and TPR. Receiver operating characteristic (ROC) curves and calibration plots were generated for the developed models.

All the statistical analyses were conducted using "The R Project for Statistical Computing" (version 4.2.3; R Core Team, GNU GPL v2, r-project.org). We used the

Scikit-Learn package (version 1.3; www.scikit-learn.org), the Anaconda Python programming environment (Conda version 23.10.0; www.anaconda.com; Python version 3.10.0; www.python.org), and Jupyter Notebooks (jupyter.org) for data processing and model building.

Results

Between January 2018 and January 2024, 6087 patients with acute injuries to at least two anatomical areas were analyzed. As shown in Figs. 1 and 5648 patients were excluded for various reasons. A total of 439 patients who met the inclusion criteria were included in the study.

The general characteristics of the sample are presented in Tables 1 and 2. There were 128 women (29.2%) and 311 men (70.8%). In 230 (52.4%) patients, SIRS was diagnosed in the first 24 h of stay, which was confirmed by the prolonged (>4 h) presence of two or more criteria: pulse (84.7 ± 14.6 vs. 100.1 ± 17.5 beats per minute; $P < 0.0001$), respiratory rate (18.7 ± 2.1 vs. 20.2 ± 3.1 breaths per minute; $P < 0.0001$), white blood cell count (14 ± 6.8 vs. $18.2 \pm 6.5 \times 10^9/L$; $P < 0.0001$), and fever (12.8% vs. 87.2% of patients; OR 11.05, 95% CI: 5.92–22.34; $P < 0.0001$).

Table 1 Patient characteristics of the non-SIRS and SIRS groups (continuous values)

	Non-SIRS = 209	SIRS = 230	Total = 439	Difference in means [95% CI]	P value
Age (years)	47.2 (15.3)	41.9 (15.1)	44.4 (15.4)	-5.4 [-8.2 to -2.6]	0.0003
NISS	26 (9)	29.8 (11.4)	28 (10.5)	3.8 [1.8 to 5.8]	0.0003
Pulse (min)	84.7 (14.6)	100.1 (17.5)	92.8 (17.9)	15.4 [12.2 to 18.6]	< 0.0001
SAP (mmHg)	114.4 (22.3)	101.4 (26.6)	107.6 (25.5)	-13.0 [-17.7 to -8.4]	< 0.0001
DAP (mmHg)	71.4 (15.2)	63.4 (17.8)	67.2 (17.1)	-8.0 [-11.0 to -5.1]	< 0.0001
Respiration Rate (min)	18.7 (2.1)	20.2 (3.1)	19.5 (2.8)	1.5 [1.0 to 1.9]	< 0.0001
RBC ($\times 10^{12}/L$)	4.2 (0.7)	4.1 (0.8)	4.1 (0.7)	-0.1 [-0.2 to 0.04]	0.3129
Hematocrit (%)	36.2 (6.1)	35.5 (6.9)	35.8 (6.5)	-0.8 [-2.0 to 0.5]	0.3601
Hemoglobin (g/L)	124.3 (21.6)	121.9 (24.7)	123 (23.3)	-2.5 [-6.7 to 1.8]	0.5402
MCH (pg)	29.9 (3.1)	30.1 (3.2)	30 (3.2)	0.2 [-0.4 to 0.8]	0.6961
MCHC (g/L)	343 (15.5)	343.4 (20.8)	343.2 (18.5)	0.5 [-2.8 to 3.7]	0.8127
MCV (fL)	87.1 (8.1)	87.5 (7.5)	87.3 (7.8)	0.4 [-1.2 to 1.9]	0.8168
WBC ($\times 10^9/L$)	14 (6.8)	18.2 (6.5)	16.2 (7)	4.2 [2.9 to 5.4]	< 0.0001
Platelets ($\times 10^9/L$)	239 (85.5)	244.3 (79.1)	241.8 (82.2)	5.3 [-9.9 to 20.5]	0.1906
Total protein (g/L)	64.2 (8.6)	61.4 (8.7)	62.7 (8.8)	-2.8 [-4.4 to -1.2]	0.0010
Total bilirubin (mcmol/L)	13.8 (8.9)	12.1 (7.5)	12.9 (8.3)	-1.8 [-3.2 to -0.4]	0.0081
Glucose (mmol/L)	8 (2.9)	9 (3.7)	8.5 (3.4)	1.0 [0.4 to 1.6]	0.0018
Creatinine (mcmol/L)	83.7 (29.9)	88.6 (27.7)	86.3 (28.8)	4.9 [-0.4 to 10.2]	0.0136
Urea (mmol/L)	5.5 (2.4)	5.5 (2.3)	5.5 (2.4)	0.007 [-0.4 to 0.5]	0.8277
RBC (units)	0.7 (1.5)	1.4 (2.3)	1.1 (2)	0.7 [0.4 to 1.1]	0.0001
Plasma (units)	0.4 (1.6)	1.2 (2.6)	0.9 (2.2)	0.8 [0.4 to 1.2]	0.0001
Fibrinogen (g/L)	2.8 (1.1)	2.5 (1.1)	2.6 (1.1)	-0.3 [-0.5 to -0.1]	0.0022
INR	1.1 (0.2)	1.2 (0.3)	1.1 (0.2)	0.1 [0.02 to 0.1]	0.0324
APTT (s)	27.6 (5)	28.4 (7.7)	28 (6.6)	0.8 [-0.4 to 2.0]	0.9229
First total care (day)	4.9 (5.2)	4.3 (5.1)	4.6 (5.2)	-0.6 [-1.5 to 0.3]	0.0540
LOS in ICU (days)	2.7 (4.3)	5 (6.3)	3.9 (5.6)	2.3 [1.3 to 3.2]	< 0.0001
LOS in Hospital (days)	16 (8.1)	16.9 (10.9)	16.5 (9.7)	0.9 [-0.9 to 2.7]	0.6792

The data are presented as the mean (standard deviation); SAP, systolic arterial pressure; DAP, diastolic arterial pressure; LOS, length of stay; ICU, intensive care unit; INR, international normalized ratio; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; APTT, activated partial thromboplastin time; NISS, new injury severity score

Table 2 Patient characteristics of the non-SIRS and SIRS groups (dichotomous values)

	Indicator	Non-SIRS = 230	SIRS = 230	Total = 439	Odds ratio [95% CI]	P value
Gender	Female	67 (52.3%)	61 (47.7%)	128 (29.2%)	1.31 [0.87 to 1.98]	0.2025
	Male	142 (45.7%)	169 (54.3%)	311 (70.8%)		
Fever	No	193 (61.5%)	121 (38.5%)	314 (71.5%)	10.87 [6.30 to 19.89]	< 0.0001
	Yes	16 (12.8%)	109 (87.2%)	125 (28.5%)		
Needs ICU	No	92 (71.9%)	36 (28.1%)	128 (29.2%)	4.24 [2.73 to 6.70]	< 0.0001
	Yes	117 (37.6%)	194 (62.4%)	311 (70.8%)		
Prolonged shock	No	144 (61.8%)	89 (38.2%)	233 (53.1%)	3.51 [2.37 to 5.23]	< 0.0001
	Yes	65 (31.6%)	141 (68.4%)	206 (46.9%)		
Hemorrhage > 25% of BV	No	132 (54.1%)	112 (45.9%)	244 (55.6%)	1.81 [1.23 to 2.65]	0.0023
	Yes	77 (39.5%)	118 (60.5%)	195 (44.4%)		
ATC	No	171 (53.3%)	150 (46.7%)	321 (73.1%)	2.40 [1.55 to 3.77]	0.0001
	Yes	38 (32.2%)	80 (67.8%)	118 (26.9%)		
Pneumonia	No	124 (54.1%)	105 (45.9%)	229 (52.2%)	1.74 [1.19 to 2.54]	0.0042
	Yes	85 (40.5%)	125 (59.5%)	210 (47.8%)		
Outcome	Deceased	22 (35.5%)	40 (64.5%)	62 (14.1%)	0.56 [0.32 to 0.97]	0.0391
	Alive	187 (49.6%)	190 (50.4%)	377 (85.9%)		

ICU, intensive care unit; BV, blood volume; ATC, acute traumatic coagulopathy

SIRS was slightly more common among 169 men (54.3%, OR=1.31, 95% CI: 0.87–1.98; $P=0.2025$). The average age of the injured patients was 44.4 ± 15.4 years; in the non-SIRS group, the average age of the patients was younger (47.2 ± 15.3 vs. 41.9 ± 15.1 ; $P=0.0003$).

In total, 3835 injuries have been reported. The most frequently damaged anatomical areas according to the AIS were the lower extremities ($n=891$; $\mu=2.01 \pm 1.96$ injuries per patient), chest ($n=846$; $\mu=1.92 \pm 1.58$), head ($n=786$; $\mu=1.79 \pm 1.84$), upper extremities ($n=412$; $\mu=0.94 \pm 1.22$), and spine ($n=338$; $\mu=0.77 \pm 1.22$). The severity of the injuries ranged from minor ($n=1149$) to fatal ($n=3$), and the average NISS value was 28.0 ± 10.5 (min=17; max=75). In the non-SIRS group, the average severity of damage was lower (26 ± 9 vs. 29.8 ± 11.4 , $P=0.0003$). The most common causes of injury were traffic accidents (total: 239 (54.4%) cases; driver: 59 (13.4%); passenger: 70 (15.9%); pedestrian: 92 (21.0%); others: 18 (4.1%)), falls (total: 151 (34.4%) cases; height ≤ 2 m: 14 (3.2%); height > 2 m: 137 (31.2%)).

When comparing non-SIRS and SIRS samples with respect to noncriteria clinical parameters, the SIRS group had lower hemodynamic indicators (pulse, arterial pressure; $P<0.0001$) upon admission to the hospital, which was associated with more frequent class II bleeding (118 (60.5%) vs. 77 (39.5%) patients; OR 1.81, 95% CI: 1.23–2.65; $P=0.0023$) and the need for long-term anti-shock therapy with vasoactive drugs (141 (68.4%) vs. 65 (31.6%) patients; OR 5.51, 95% CI: 2.37–5.23; $P<0.0001$). In the same patients, the average need for erythrocyte transfusion in the first 48 h after admission was almost twice as high, and for plasma, it was three times greater ($P<0.0001$), as was the risk of developing OTC (80 (67.8%) vs. 38 (32.2%) cases; OR 2.4, 95% CI: 1.55–3.77;

$P<0.0001$). More severe bleeding in the SIRS group was associated with hypoproteinemia (61.4 ± 8.7 vs. 64.2 ± 8.6 g/L; $P=0.001$), hypofibrinogenemia (2.5 ± 1.1 vs. 2.8 ± 1.1 g/L; $P=0.0022$) and increased INR (1.2 ± 0.3 vs. 1.1 ± 0.2 ; $P=0.0324$).

At the first biochemical blood test, there were statistically significant differences in the levels of total bilirubin and creatinine; however, they were within normal values. Glucose levels were greater in the SIRS group, which may be related to the severity of injuries and pain syndrome against this background or to the use of adrenergic drugs in anti-shock therapy (9 ± 3.7 vs. 8 ± 2.9 mmol/L; $P<0.0001$).

Patients with SIRS were at greater risk of developing pneumonia (125 (59.5%) vs. 85 (40.5%) patients; OR 1.74, 95% CI: 1.19–2.54; $P=0.0042$). In patients with SIRS, the need for prior ICU admission was almost two times greater they had a longer ICU LOS (5 ± 6.3 vs. 2.7 ± 4.3 days; $P<0.0001$). There were no statistically significant differences in the timing of definitive surgery; on average, all patients underwent delayed surgery on 4.6 ± 5.2 days of hospital stay. In 59 (13.4%) patients, only conservative treatment was performed, 15 of whom died without surgical care due to the severity of their condition. Temporary external fixation followed by delayed definitive surgery was performed in 183 patients (41.7%). Primary definitive surgery was performed in 177 patients (40.3%) at a mean of nine days after hospitalization. Surgical tactics for the abdominal organs were applied in the remaining 64 (14.6%) patients, followed by conservative trauma management. The total duration of hospitalization did not significantly differ ($\mu=16.6 \pm 9.7$ days; IQR=11–21); however, patients with SIRS had a greater number of

adverse outcomes (40 (64.5%) vs. 22 (35.5%) deaths; OR 10.87, 95% CI: 6.3–19.89; $P=0.0391$).

A total of 439 patients with severe polytrauma were divided into three datasets. First, the external validation dataset ($n=109$ (24.8%) patients) was randomly generated. The remaining patients were used for model development. This subset was divided into training ($n=231$ (52.6%) patients) and internal validation ($n=99$ (22.6%) patients) datasets. The general characteristics of the model development and validation datasets are presented in Supplementary Tables 1 and 2, respectively. There was no statistically significant difference between the indicators except for the NISS score, for which the average injury severity in the development dataset was slightly greater (28.6 ± 10.7 vs. 26.3 ± 9.8 points; $P=0.0149$).

Model development and internal validation

To predict the risk of SIRS development within 24 h of admission to the emergency department, six models were developed: the DTC, RFC, SVC, LR, DTC, and BBC models. In the internal validation stage, all models were effective in predicting the early development of SIRS (AUROC for all models ≥ 0.87) (Fig. 2). In particular, in the scoring system, the BBC showed the best result, with 40 points (AUROC 0.87; 95% CI: 0.80–0.94). The model had the overall FPR of 0.09 with a relatively high TPR of 0.77, and the total accuracy was the highest among all the developed models (0.84). RFC was the second most effective model, with a total score of 34, which had the highest

AUROC (0.89; 95% CI: 0.83–0.96), while the FPR (0.13) and TPR (0.77) were quite close to those of the BBC. The accuracy of RFC was the second-best indicator (0.82) among the models. The other four models had problems with the overall FPR and/or TPR while maintaining their accuracy at 0.79. A detailed representation of the effectiveness of the developed models during the internal validation stage is presented in Supplementary Table 3. To further analyze the ability of the models to predict the risk of developing SIRS, density curves (Fig. 3) were generated, which represent the distribution of predicted values and their relation to the decision threshold. All models showed a relatively small overlap and a large differentiation area, indicating better discrimination. To evaluate whether the probabilistic outputs can be interpreted as the probability of an event, all the developed models were calibrated using the sigmoid cross-validation method. In the calibration graphs (Supplementary Fig. 1), the best predictive ability was characterized by the location of the model curve closest to the ideal calibration line over the entire range of empirical and predicted capabilities. Deviations from the ideal line indicate either excessive or insufficient confidence in model forecasts. Three models (RFC, SVC, and LR) showed the best predictive ability; however, there were deviations in some ranges for each of these models. When calibrating all the developed models using logistic regression, the prediction results of all the models tended to the ideal values and did not differ from each other.

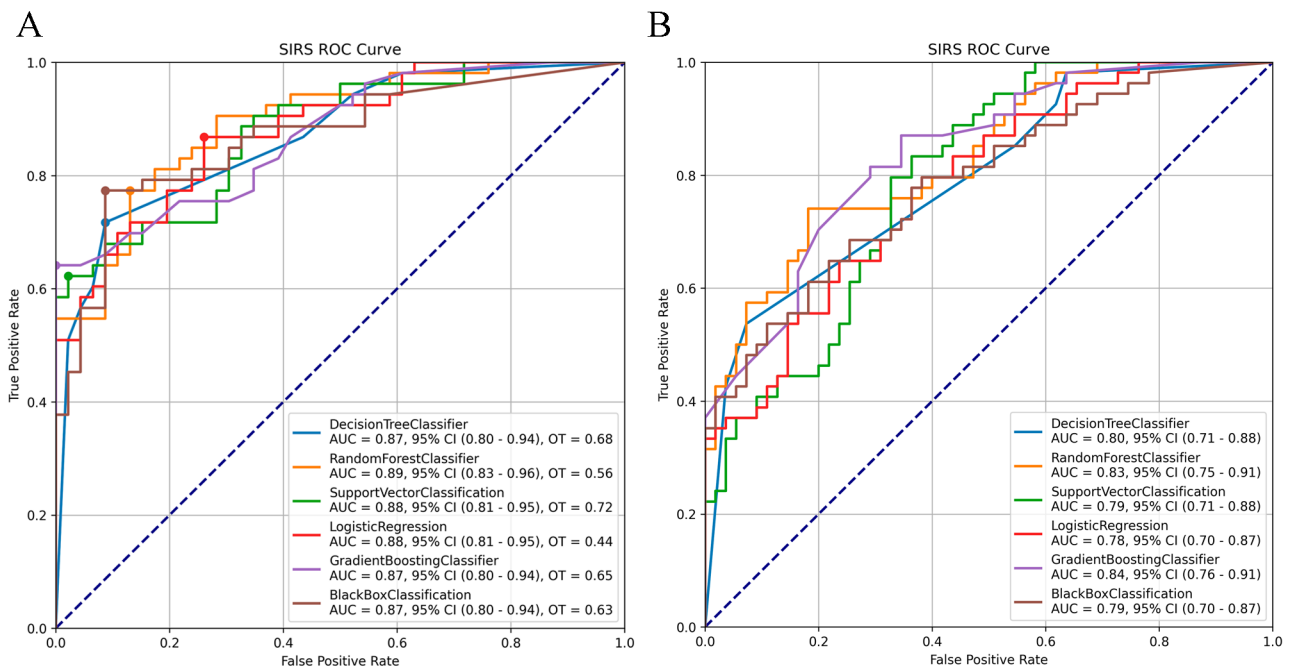


Fig. 2 Receiver operating characteristic curves of all models. SIRS, systemic inflammatory response syndrome; ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval; OT, optimal threshold. **A:** ROC curve of all models based on the internal validation; **B:** ROC curve of all models based on the external validation

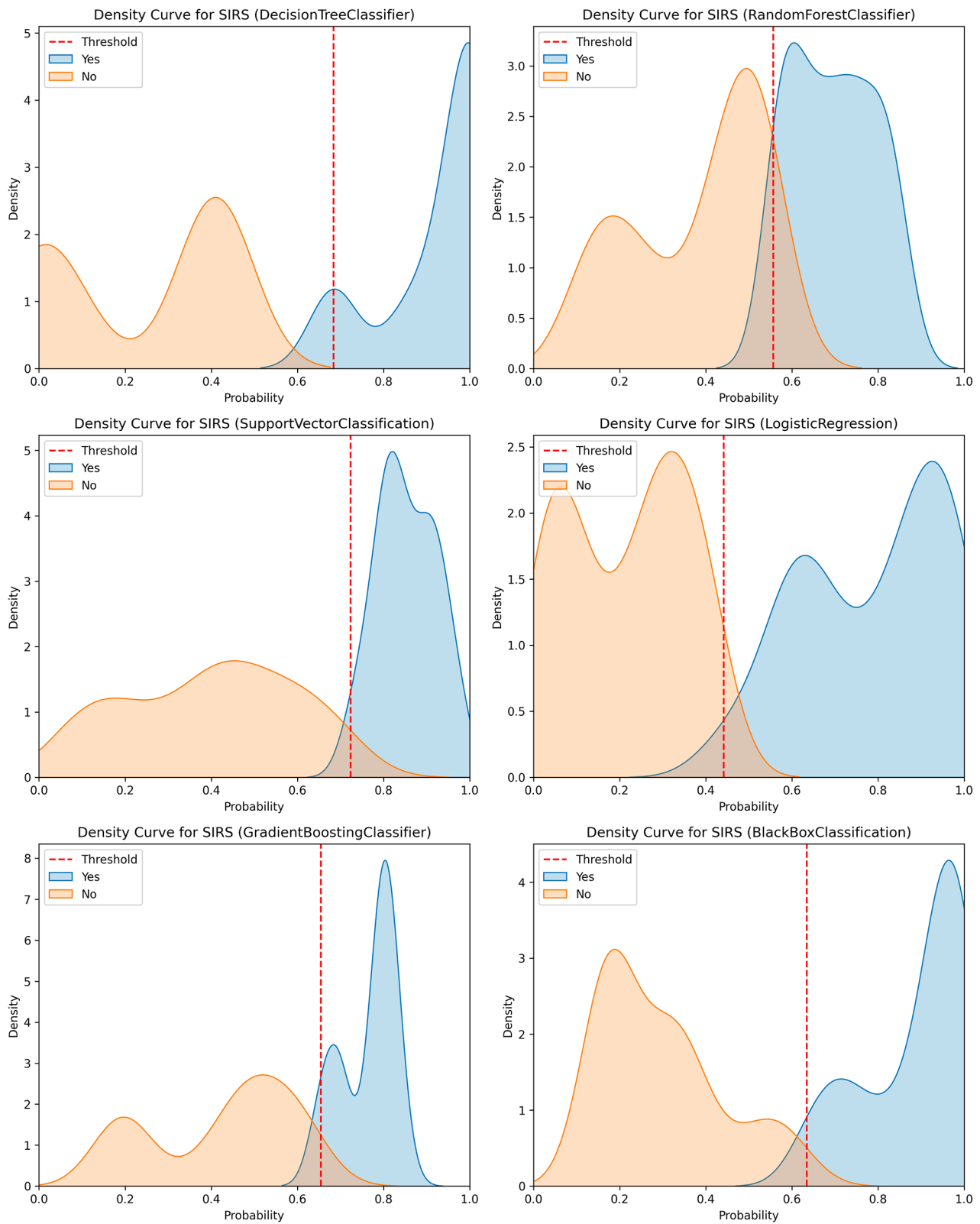


Fig. 3 Density curves for all models. Density curves represent the distribution of predicted values and their relation to the optimal decision threshold. Orange indicates patients without SIRS, and blue indicates patients with SIRS. The smaller the overlap between the colors, the better the model's ability to discriminate

External model validation

External validation was performed by using noncalibrated models to evaluate the best basic predictive ability. The efficiency of all models decreased as expected (AUROC for all models ≥ 0.78). In the scoring system for the validation dataset, the best result was noted in the RFC model, with 47 points (AUROC 0.83; 95% CI: 0.75–0.91), showing a relative improvement in results compared to internal validation. Two models had 40 points: GBC (AUROC 0.84; 95% CI: 0.76–0.91) and DTC (AUROC 0.80; 95% CI: 0.71–0.88). In turn, BBC showed a significant deterioration in results, taking third place with 30 points (AUROC 0.79; 95% CI: 0.70–0.87) due to a deterioration in the overall FPR (0.22) and TPR (0.61). The remaining three models had problems with a deterioration in the initial values for the overall FPR and/or TPR while maintaining their accuracy at 0.65 (Supplementary Table 4).

To determine the best model based on the results of internal and external validation, the score tables were summed in a ratio of 0.33 and 0.67, respectively. According to the final model rating table (Table 3), RFC was recognized as the most stable model in terms of efficiency, scoring 42.7 points. The precision of recognition for non-SIRS patients was 0.76, whereas for SIRS patients, it was 0.78. The ability to determine the non-SIRS class was characterized by a recall of 0.80 and 0.74 for the SIRS group. The percentage of correct answers for the model was represented by an accuracy value of 0.77, which was higher than that of the other models in the external validation stage. The RFC model had one of the best overall TPR indicators (0.74) and one of the lowest FPR indicators (0.20).

Using the SHAP method, data on the importance of features were extracted from the RFC. Because the RFC model consisted of 56 decision trees, we used the top 15

features to further study their impact on SIRS. The features included the NISS; weighted AIS severity for head, abdomen, and lower extremity injuries; leukocyte count and excess; fever; hemoglobin; MCV; MCH; MCHC; and serum total protein, glucose, urea, and fibrinogen (Fig. 4).

Discussion

Main findings

This study is part of another project to develop comprehensive software for patients with severe polytrauma. Owing to the large amount of information and data, the current work presents the results of studying one of the endpoints – SIRS.

In this study, we developed and validated machine learning-based models capable of predicting the possible development of SIRS in the next 24 h, even at the admission department stage, using a minimal set of diagnostic tests. Of the six models tested, the RFC model was determined to be optimal because it maintained a high level of discrimination during internal and external validation. Using a minimal set of examinations conducted in the admission department, our model allowed us to predict the risk of developing SIRS with high accuracy.

While analyzing why the random forest classifier had greater success than the other models, the following opinions were made. The working principle of linear regression and support vector machines is based on building a delimiting hyperplane in an n-dimensional data array, where n is the number of features used. Both methods use each feature once, which implies the presence of a certain “rigid” mathematical relationship between the features, leading to specific results. With respect to polytrauma, the presence of such dependencies is possible; however, when solving the problem of predicting the development of SIRS, we did not find sufficient evidence for such relationships (SVC was in sixth

Table 3 Ranking of the prediction performance of all models based on both internal and external validation for SIRS using a scoring system

	Decision Tree Classifier	Random Forest Classifier	Support Vector Classification	Logistic Regression	Gradient Boosting Classifier	Black Box Classification
0 Precision	3.00	4.00	1.00	4.66	2.00	3.66
0 Recall	3.67	2.67	5.34	1.00	5.00	2.33
0 F1-Score	4.34	4.34	2.67	1.00	4.00	2.99
1 Precision	3.67	2.67	5.67	1.00	5.33	2.66
1 Recall	3.00	4.67	1.00	5.67	2.00	4.00
1 F1-Score	3.00	4.67	1.00	4.33	2.00	3.99
Accuracy	3.34	4.34	1.00	2.00	2.67	3.33
FPR	3.67	2.67	5.34	1.00	5.00	2.33
TPR	3.00	4.67	1.00	5.67	2.00	4.00
AUROC	2.34	3.67	2.00	1.33	3.68	1.67
AUPRC	3.01	4.34	1.33	2.00	4.01	2.34
Total points	36.04	42.71	27.35	29.66	37.69	33.3

FPR, false-positive rate; TPR, true-positive rate; AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve

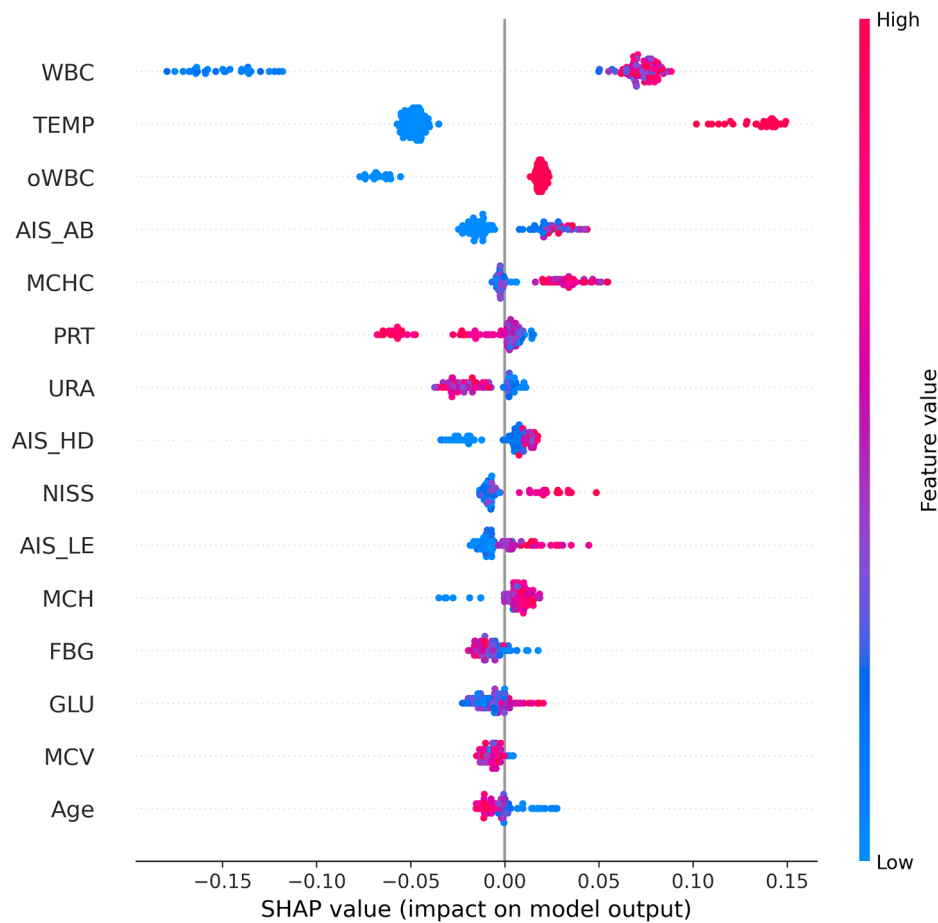


Fig. 4 Feature importance analysis using the SHapley Additive exPlanation (SHAP). SHAP feature importance analysis of the internal validation cohort. The X-axis of the graph represents the impact of the feature on the prediction result. The Y-axis represents the model predictors. The higher the feature on the graph, the stronger is the correlation between the feature and the prediction result. Blue indicates a low feature value, whereas pink indicates a high feature value. The top 15 features are as follows: WBC, leukocyte count; TEMP, fever; oWBC, excess leukocyte count; AIS_AB, weighted AIS severity for abdominal injuries; MCHC, mean corpuscular hemoglobin concentration; PRT, serum total protein; URA, serum urea; AIS_HD, weighted AIS severity for head injuries; NISS, new injury severity score; AIS_LE, weighted AIS severity for lower extremity injuries; MCH, mean corpuscular hemoglobin; FBG, fibrinogen; GLU, serum glucose; and MCV, mean corpuscular volume

place in the scoring system, with 27.35 points, and LR was in fifth place, with 29.66 points).

The results of a neural network depend mostly on its architecture. In our case, we used a deep feed-forward (DFF) architecture with automatic tuning of parameters, including the number of hidden layers and nodes. The final model contained two hidden layers with 18 and 9 nodes. Considering the logic of the neural networks, each feature has a certain weight that affects the final result. The value of the node from the first layer affects each node in the hidden layer, which is essentially a combination of all previous features. Thus, the initial features can be reused indirectly. However, high-quality training of the model is required to achieve optimal results. On the basis of the results of our study (fourth place in the scoring system with 33.30 points), the number of cases for training was probably insufficient.

The method of constructing a logical tree can be considered an intermediate method between neural networks and linear algorithms. A decision tree constructs a regression tree on the basis of binary recursive partitioning, iteratively splitting the data into partitions until each node reaches a user-specified minimum node size. The splitting process is applied to each new branch and continues until the entire structure is reached. The positive aspect is that the decision tree can use highly correlated features repeatedly, creating alternative paths leading to an answer. In our opinion, the result of the decision tree in the study (third place in the scoring system with 36.04 points) is due to the complexity of the problem, which does not imply the presence of just one simple tree.

On the basis of this conclusion, the use of multiple weak learners and decision trees may be an alternative solution. Boosting is an ensemble learning method that

combines a set of weak learners into a strong learner to minimize training errors. The gradient boosting method is potentially the most accurate in prediction because each weak learner recycles the signs, creating alternative views on the final result. However, similar to neural networks, boosters work more efficiently with big data. In our study, the model had the second-best score at 37.69 points.

Unlike boosting, the random forest does not use various prediction techniques but combines multiple decision trees into a “forest”, each constructed using bootstrap samples of the training data and random feature selection. In this ensemble, each tree relies on the values of a random vector sampled independently and drawn from the same distribution for all the trees in the forest. Prediction is then achieved by aggregating individual tree predictions through methods such as majority voting or averaging, resulting in a robust and versatile model for various machine learning tasks. Multiple reuses of primary features with alternative result branches in our study yielded the best result (the first place with 42.71 points). Thus, in our opinion, the absence of restrictions on the reuse of variables in small datasets directly affects the prediction results.

Systemic inflammatory response syndrome

SIRS is one of the most common pathological processes in patients with polytrauma. According to NeSmith et al., 79% of patients with an ISS > 16 points had signs of SIRS at the time of admission to the emergency department and needed to be hospitalized in the ICU for intensive care [27]. In turn, Butcher et al. reported that in a level-1 urban trauma center, SIRS was observed in 81% of patients with severe polytrauma (ISS > 15 points) at least once in the first 72 h after hospitalization [28].

SIRS was diagnosed based on two of the four criteria [24]. However, Schefzik et al. correctly noted that with polytrauma, patients can receive a certain type of treatment that directly affects the criteria for the presence of SIRS. Thus, the use of vasopressant agents in shock therapy affects heart rate, and with artificial ventilation, it is not possible to estimate the true respiratory rate [29]. In this regard, a dynamic assessment of the patient's condition is needed, adjusted for therapy and concomitant diseases of the heart and respiratory system. When analyzing the data, we considered the specifics of therapy and anamnesis as much as possible when making a diagnosis of SIRS. To minimize errors, the continuous retention of two or more SIRS criteria for 4 h was used as a direct basis for diagnosis. Thus, in our study, the incidence of SIRS was lower than that reported by the authors (52.4%). This difference may be due to a prolonged assessment of the SIRS criteria, a shorter time window of observation

(24 h after hospitalization), or the correction of criteria taking into account the somatic status.

As the incidence of SIRS is also closely correlated with the severity of the injury, the length of stay of such patients in the ICU naturally increases. For example, according to NeSmith et al., polytrauma patients spend an average of 19.2 days in the ICU for four positive SIRS criteria, 8.2 days for three criteria and 6.3 days for two criteria. Patients in the non-SIRS group stayed for up to 4.7 days on average. In our study, for objective reasons, we did not study the distribution of ICU stays relative to the number of available SIRS criteria. Nevertheless, patients in the SIRS group stayed in the ICU for twice as long (mean 5 vs. 2.7 days). Such short periods of stay are due to the small number of beds (up to 9–12 beds per organization) and the large flow of patients. The organizational principles of the intensive care unit in the Republic of Kazakhstan imply the use of the ICU not only as an intensive care unit but also as a ward for early post-operative management of patients. After the stabilization of vital signs and in the absence of a threat to life, patients in any treatment group are transferred to the Department of Orthosurgery for further observation and treatment.

Use of machine learning models

Currently, we are witnessing active development and implementation of ML techniques in clinical practice. The large abundance of different ML models for solving the same problem is due to the heterogeneous data and conditions in which the study was conducted. Therefore, the application of a model developed based on the data of one clinic may be ineffective under the conditions of another clinic.

The problem of predicting SIRS development has gained popularity in recent years [30–32]. Mica et al. developed a visual analytical tool for assessing a patient's condition (Sankey diagram) [30]. The endpoints of the assessment were the risk of developing SIRS, sepsis, and early (<72 h) mortality. The predictive model was built on 1925 patients with severe trauma (ISS > 16) and was based on shock indicators (pH, lactate, temperature, hemoglobin, and hematocrit). The developed model visually demonstrates the risks of complications and methods that bring the greatest weight to the results. According to the authors, such models make it possible to plan the tactics of surgical management as well as to target therapy regarding the risks of developing coagulopathy, hemorrhagic shock, and treatment outcomes. Unfortunately, this tool has not been validated or evaluated for its effectiveness.

Another study on the effect of SIRS on outcomes was conducted by Fachet et al., where machine learning was used to determine the immunological patterns of the development of SIRS, pneumonia, and sepsis [31]. The

frequency of SIRS among injured patients ($ISS > 16$) at the admission department stage was 56.9%, and on the second day, it was 39.2%. The study revealed a weak correlation between immunological markers and SIRS, while there was a strong correlation between the following indicators of massive tissue damage: severity of injury according to ISS, SHG score, leukocyte level, prothrombin time, INR, APTT, hematocrit, and hemoglobin. Thus, patients with head and chest injuries are at greater risk of developing SIRS. When divided into clusters, these patients showed a proinflammatory pattern due to increased levels of IL-6 and IL-10. Unfortunately, Facht et al. did not develop a separate model for predicting SIRS development. First, the SIRS criteria overlapped with those for sepsis and pneumonia. Second, the weak correlation between SIRS incidence and immunological markers prevented them from building a final model.

In our study, the relationships between SIRS, the volume of damaged tissue, and concomitant complications were also clearly visible. In addition to the direct SIRS criteria, the best RFC model for SHAP analysis also considers a weighted assessment of the severity of injury according to the AIS for the head, abdomen, and lower extremities. Brain damage, as well as chest damage, directly affects three of the four SIRS criteria: thermoregulation [33], heart activity [34], and respiration [35, 36]. In turn, owing to the large number of parenchymal organs in the abdominal cavity, bleeding in patients with abdominal trauma is more severe, which leads to sharp shifts in red blood indicators. Our RFC model also indicated the association of MCV, MCH, and MCHC with the risk of developing SIRS in the first 24 h. The above-mentioned studies, together with our results, confirm the need to review SIRS criteria in patients with severe polytrauma.

Limitations

Our study had several limitations. Because the work done is part of another project, we were severely limited in the amount of data. All selected clinical cases had to meet the requirements of all parallel studies simultaneously. The retrospective and prospective nature of the data collection process is a limitation. Unfortunately, there is no unified registry of patients with polytrauma in the Republic of Kazakhstan, which has led to a shortage of high-quality data. Despite working in the largest EMRs system, we faced the problem of multiple defects in filling out the documentation. The EMRs were not sufficiently structured, which caused difficulties in finding and extracting the necessary indicators. Therefore, we excluded many potentially useful features from the final dataset.

Another problem is the approach used to describe the patient's injury. The specialists of the clinics adhered to

the traditional methods of describing fractures, mainly using the IDC-10 and AO Muller fracture classifications. In the Republic of Kazakhstan, injury coding using the AIS is not mandatory. Therefore, the authors had to retrospectively encode all existing injuries, following the recommendations of the AIS (update 2008), based on data from the initial medical examination, records of consulting doctors, and diagnostic and intraoperative findings. This fact served as one of the grounds for expanding the criteria for the new Berlin definition of polytrauma.

The model was developed based on data from three trauma clinics, where common approaches to early therapy could differ from each other. Despite the demonstrated predictive effectiveness, the model additionally requires external validation using more data from different samples and under different conditions. Additional studies of the aspects of the problem will contribute to improving the prediction results of the model, which will ultimately increase its clinical value. Currently, it is extremely important to test the clinical practicality of using a model for predicting the development of SIRS in the next 24 h in patients with severe polytrauma when planning and conducting treatment for subsequent assessments of changes in the timing of definitive care, ICU and hospital stays, and changes in mortality rates.

Conclusions

Machine learning allows for the determination of complex relationships or patterns based on empirical data. Early recognition of the potential development of SIRS using ML techniques will allow doctors to consider treatment tactics carefully. Unfortunately, different medical approaches to the management of patients with polytrauma, from the accident scene to the moment of discharge from the hospital, directly affect all aspects of treatment. Therefore, it is impossible to create a single generally accepted ML model that works equally effectively under different samples and conditions. To achieve similar results, it is necessary to include a large amount of data from clinics at various levels and possibly even from countries. The developed model will be integrated into a comprehensive clinical decision support system for patients with severe polytrauma.

Abbreviations

AIS	Abbreviated injury scale
APTT	Activated partial thromboplastin time
ATC	Acute traumatic coagulopathy
BBC	Black box classification
BV	Blood volume
DAP	Diastolic arterial pressure
DTC	Decision tree classifier
EMRs	Electronic medical records
GBC	Gradient boosting classifier
INR	International normalized ratio
LOS	Length of hospital stay
LR	Logistic regression

ML	Machine learning
NISS	New Injury Severity Score
RFC	Random forest classifier
SAP	Systolic arterial pressure
SHAP	SHapley Additive exPlanation
SIRS	Systemic inflammatory response syndrome
SVC	Support vector classifier

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-024-02640-x>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

Not applicable.

Author contributions

AP and MZh contributed to the idea; SA and AM were involved to the patient selection and in the data collection; AP was involved in development and validation of machine learning models; AT and MZh was involved in the conceptualization and supervision, statistical analysis; AP, AT and MZh took part in the draft writing; AP and MZh contributed to the final writing.

Funding

This study was part of the project Reg.No. AP13067824 "Development and optimization of methods for diagnosis and surgical rehabilitation of injuries using artificial intelligence and robotics" of the Ministry of Science and Higher Education of the Republic of Kazakhstan. This article presents the results of one of the final points of the project. In particular, the prediction of the development of systemic inflammatory response syndrome in patients with severe polytrauma has been studied.

Data availability

The data that support the findings of this study are available from Ministry of Healthcare of the Republic of Kazakhstan but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Ministry of Healthcare of the Republic of Kazakhstan.

Declarations

Ethics approval and consent to participate

The study received approval from the Local Ethical Committee of the Non-Commercial Joint Stock Company "Semey Medical University" in Semey city, Republic of Kazakhstan (permission No. 2, dated 10/28/2020). Additionally, it was retrospectively registered in the ClinicalTrials.gov database of the National Library of Medicine (<https://www.nlm.nih.gov/>) under the identifier NCT06323096. The need for consent to participate was waived by an Institutional Review Boards of participating centers «Municipal state enterprise with the right of economic management Emergency hospital of the Abai region» (Semey, the Republic of Kazakhstan, act No. act No. 06–01/366, dated 02/14/2023); «Multidisciplinary Hospital named after Prof. H. Makazhanov» (Karaganda, the Republic of Kazakhstan, act No. 06–01/367, dated 02/14/2023); «The National Scientific Center of Traumatology and Orthopaedics named after Academician Batpenov N.D.» (Astana, the Republic of Kazakhstan, act No. 06–01/368, dated 02/14/2023).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹University Hospital of Non-Commercial Joint-Stock Company "Semey Medical University", 1a, Ivan Sechenov str, Semey city 071400, Republic of Kazakhstan

²Center of habilitation and rehabilitation of persons with disabilities of the Abai region, 109, Karagaily, Semey city 071400, Republic of Kazakhstan

³Non-Commercial Joint-Stock Company "Semey Medical University", 103, Abai Kunanbayev str, Semey city 071400, Republic of Kazakhstan

Received: 9 May 2024 / Accepted: 20 August 2024

Published online: 27 August 2024

References

1. Gebhard F, Huber-Lang M. Polytrauma - Pathophysiology and management principles. *Langenbeck's Arch Surg.* 2008;393:825–31. <https://doi.org/10.1007/s00423-008-0334-2>.
2. Nicola R. Early total care versus damage control: current concepts in the Orthopedic Care of Polytrauma patients. *ISRN Orthop.* 2013;2013:1–9. <https://doi.org/10.1155/2013/329452>.
3. Halvachizadeh S, Baradaran L, Cinelli P, et al. How to detect a polytrauma patient at risk of complications: a validation and database analysis of four published scales. *PLoS ONE.* 2020;15:1–16. <https://doi.org/10.1371/journal.pone.0228082>.
4. Pape H-C, Tscherne H. Early definitive fracture fixation with Polytrauma: advantages Versus Systemic/Pulmonary consequences. Multiple organ failure. New York, NY: Springer New York; 2000. pp. 279–90.
5. Pfeifer R, Teuben M, Andruszkow H, et al. Mortality patterns in patients with multiple trauma: a systematic review of autopsy studies. *PLoS ONE.* 2016;11:1–9. <https://doi.org/10.1371/journal.pone.0148844>.
6. Fröhlich M, Lefering R, Probst C, et al. Epidemiology and risk factors of multiple-organ failure after multiple trauma: an analysis of 31,154 patients from the TraumaRegister DGU. *J Trauma Acute Care Surg.* 2014;76:921–7. <https://doi.org/10.1097/TA.0000000000000199>.
7. Cimbanassi S, O'Toole R, Maegele M et al. (2020) Orthopedic injuries in patients with multiple injuries: Results of the 11th trauma update international consensus conference Milan, December 11, 2017.
8. Rau CS, Wu SC, Kuo PJ, et al. Polytrauma defined by the new berlin definition: a validation test based on propensity-score matching approach. *Int J Environ Res Public Health.* 2017;14:4–13. <https://doi.org/10.3390/ijerph14091045>.
9. Carlino W. Damage control resuscitation from major haemorrhage in polytrauma. *Eur J Orthop Surg Traumatol.* 2014;24:137–41. <https://doi.org/10.1007/s00590-013-1172-7>.
10. Schwing L, Faulkner TD, Bucaro P, et al. Trauma Team activation: accuracy of Triage when Minutes Count: a synthesis of literature and performance improvement process. *J Trauma Nurs.* 2019;26:208–14. <https://doi.org/10.1097/JTN.0000000000000450>.
11. Cole E, Gillespie S, Vulliamy P, et al. Multiple organ dysfunction after trauma. *Br J Surg.* 2020;107:402–12. <https://doi.org/10.1002/bjs.11361>.
12. Baek JH, Kim MS, Lee JC, Lee JH. Systemic inflammation response syndrome score predicts the mortality in multiple trauma patients. *Korean J Thorac Cardiovasc Surg.* 2014;47:523–8. <https://doi.org/10.5090/kjtc.2014.47.6.523>.
13. Bochicchio GV, Napolitano LM, Joshi M, et al. Systemic inflammatory response syndrome score at admission independently predicts infection in blunt trauma patients. *J Trauma - Inj Infect Crit Care.* 2001;50:817–20. <https://doi.org/10.1097/00005373-200105000-00007>.
14. Kuhne CA, Ruchholtz S, Kaiser GM, et al. Mortality in severely injured elderly trauma patients—when does age become a risk factor? *World J Surg.* 2005;29:1476–82. <https://doi.org/10.1007/S00268-005-7796-Y>.
15. Kocuvan S, Brilej D, Stropnik D, et al. Evaluation of major trauma in elderly patients - a single trauma center analysis. *Wien Klin Wochenschr.* 2016;128:535–42. <https://doi.org/10.1007/S00508-016-1140-4>.
16. Mörs K, Wagner N, Sturm R, et al. Enhanced pro-inflammatory response and higher mortality rates in geriatric trauma patients. *Eur J Trauma Emerg Surg.* 2021;47:1065–72. <https://doi.org/10.1007/s00068-019-01284-1>.
17. Yourc'h M, Roquilly A, Asehnoune K. Trauma-induced damage-associated molecular patterns-mediated remote organ injury and immunosuppression in the acutely ill patient. *Front Immunol.* 2018;9. <https://doi.org/10.3389/fimmu.2018.01330>.

18. Scherer J, Coimbra R, Mariani D, et al. Standards of fracture care in polytrauma: results of a Europe-wide survey by the ESTES polytrauma section. *Eur J Trauma Emerg Surg*. 2022. <https://doi.org/10.1007/s00068-022-02126-3>.
19. Stonko DP, Guillaumondegui OD, Fischer PE, Dennis BM. Artificial intelligence in trauma systems. *Surg (United States)*. 2021;169:1295–9. <https://doi.org/10.1016/j.surg.2020.07.038>.
20. Choi A, Choi SY, Chung K, et al. Development of a machine learning-based clinical decision support system to predict clinical deterioration in patients visiting the emergency department. *Sci Rep*. 2023;13:1–10. <https://doi.org/10.1038/s41598-023-35617-3>.
21. Ehrlich H, McKenney M, Elkbuli A. The niche of artificial intelligence in trauma and emergency medicine. *Am J Emerg Med*. 2021;45:669–70. <https://doi.org/10.1016/j.ajem.2020.10.050>.
22. Pape HC, Lefering R, Butcher N, et al. The definition of polytrauma revisited: an international consensus process and proposal of the New Berlin definition. *J Trauma Acute Care Surg*. 2014;77:780–6. <https://doi.org/10.1097/TA.0000000000000453>.
23. Osler T, Baker SP, Long W. A modification of the Injury Severity score that both improves accuracy and simplifies Scoring. *J Trauma Inj Infect Crit Care*. 1997;43:922–6. <https://doi.org/10.1097/00005373-199712000-00009>.
24. Chakraborty RK, Burns B. (2023) Systemic Inflammatory Response Syndrome.
25. Gennarelli TA, Wodzin E. others (2008) Abbreviated injury scale 2005: update 2008. *Russ Reeder* 200:2008.
26. Shi X, Cui Y, Wang S, et al. Development and validation of a web-based artificial intelligence prediction model to assess massive intraoperative blood loss for metastatic spinal disease using machine learning techniques. *Spine J*. 2024;24:146–60. <https://doi.org/10.1016/j.spinee.2023.09.001>.
27. NeSmith EG, Weinrich SP, Andrews JO, et al. Systemic inflammatory response syndrome score and race as predictors of length of stay in the intensive care unit. *Am J Crit Care*. 2009;18:339–46. <https://doi.org/10.4037/ajcc2009267>.
28. Butcher NE, Balogh ZJ. The practicality of including the systemic inflammatory response syndrome in the definition of polytrauma: experience of a level one trauma centre. *Injury*. 2013;44:12–7. <https://doi.org/10.1016/j.injury.2012.04.019>.
29. Schefzik R, Hahn B, Schneider-Lindner V. Dissecting contributions of individual systemic inflammatory response syndrome criteria from a prospective algorithm to the prediction and diagnosis of sepsis in a polytrauma cohort. *Front Med*. 2023;10. <https://doi.org/10.3389/fmed.2023.1227031>.
30. Mica L, Niggli C, Bak P, et al. Development of a Visual Analytics Tool for Polytrauma patients: Proof of Concept for a New Assessment Tool using a multiple layer Sankey Diagram in a single-center database. *World J Surg*. 2020;44:764–72. <https://doi.org/10.1007/s00268-019-05267-6>.
31. Fachel M, Mushunuri RV, Bergmann CB, et al. Utilizing predictive machine-learning modelling unveils feature-based risk assessment system for hyperinflammatory patterns and infectious outcomes in polytrauma. *Front Immunol*. 2023;14:1–15. <https://doi.org/10.3389/fimmu.2023.1281674>.
32. Li X, Lu Y, Chen C, et al. Development and validation of a patient-specific model to predict postoperative SIRS in older patients: a two-center study. *Front Public Heal*. 2023;11:1–10. <https://doi.org/10.3389/fpubh.2023.1145013>.
33. Thompson HJ, Tkacs NC, Saatman KE, et al. Hyperthermia following traumatic brain injury: a critical evaluation. *Neurobiol Dis*. 2003;12:163–73. [https://doi.org/10.1016/S0969-9961\(02\)00030-X](https://doi.org/10.1016/S0969-9961(02)00030-X).
34. Keane RW, Hadad R, Scott XO, et al. Neural–cardiac Inflammation Axis after Traumatic Brain Injury. *Pharmaceuticals*. 2023;16:1382. <https://doi.org/10.3390/ph16101382>.
35. Ma RN, He YX, Bai FP, et al. Machine Learning Model for Predicting Acute Respiratory failure in individuals with moderate-to-severe traumatic brain injury. *Front Med*. 2021;8. <https://doi.org/10.3389/fmed.2021.793230>.
36. Kerr N, de Rivero Vaccari JP, Dietrich WD, Keane RW. Neural-respiratory inflammation axis in traumatic brain injury. *Exp Neurol*. 2020;323:113080. <https://doi.org/10.1016/j.expneurol.2019.113080>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.