Psychiatria Danubina, 2023; Vol. 35, No. 1, pp 62-72 https://doi.org/10.24869/psyd.2023.62 *Original paper*

INTERPRETABLE ESTIMATION OF SUICIDE RISK AND SEVERITY FROM COMPLETE BLOOD COUNT PARAMETERS WITH EXPLAINABLE ARTIFICIAL INTELLIGENCE METHODS

Neslihan Cansel[1], Fatma Hilal Yagin[2], Mustafa Akan[3] & Bedriye Ilkay Aygul[1]
[1]*Inonu University Faculty of Medicine, Department of Psychiatry, Malatya, Turkey*
[2]*Inonu University Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Turkey*
[3]*Malatya Training and Research Hospital, Department of Psychiatry, Malatya, Turkey*

received: 14.6.2022; revised: 30.8.2022; accepted: 20.9.2022

SUMMARY

Background: *The peripheral inflammatory markers are important in the pathophysiology of suicidal behavior. However, methods for practical uses haven't been developed enough yet. This study developed predictive models based on explainable artificial intelligence (xAI) that use the relationship between complete blood count (CBC) values and suicide risk and severity of suicide attempt.*

Subjects and methods: *544 patients who attempted an incomplete suicide between 2010-2020 and 458 healthy individuals were selected. The data were obtained from the electronic registration systems. To develop prediction models using CBC values, the data were grouped in two different ways as suicidal/healthy and attempted/non-attempted violent suicide. The data sets were balanced for the reliability of the results of the machine learning (ML) models. Then, the data was divided into two; 80% of as the training set and 20% as the test set. For suicide prediction, models were created with Random Forest, Logistic Regression, Support vector machines and XGBoost algorithms. SHAP, was used to explain the optimal model.*

Results: *Of the four ML methods applied to CBC data, the best-performing model for predicting both suicide risk and suicide severity was the XGBoost model. This model predicted suicidal behavior with an accuracy of 0.83 (0.78-0.88) and the severity of suicide attempt with an accuracy of 0.943 (0.91-0.976). Lower levels of NEU, WBC, MO, NLR, MLR and, age higher levels of HCT, PLR, PLT, HGB, RBC, EO, MPV and, BA contributed positively to the predictive created model for suicide risk, while lower PLT, BA, PLR and RBC levels and higher MO, EO, HCT, LY, MLR, NEU, NLR, WBC, HGB and, age levels have a positive contribution to the predictive created model for violent suicide attempt.*

Conclusion: *Our study suggests that the xAI model developed using CBC values may be useful in detecting the risk and severity of suicide in the clinic.*

Key words: *suicidality - violent suicide - nonviolent suicide – CBC - machine learning - Explainable Artificial Intelligence*

* * * * *

INTRODUCTION

Suicide, often as a result of mental illness, is the act of taking one's own life intentionally (Bhatia et al. 2000). According to the World Health Organization, approximately 700,000 people die each year by suicide. It is estimated that there are between 10 and 25 non-fatal suicide attempts to complete each suicide, increasing to 100-200 for adolescents. Suicide is the fourth leading cause of death between the ages of 15-29, and as such it is an important public health problem (Maris 2002, WHO 2021).

Although suicide is seen in almost every society, every economic level and gender, it is generally seen in men, in low and middle economic level countries, in the presence of comorbid psychiatric disease (essentially depression and alcoholism), and is more common in those who have attempted suicide before, in those with a family history of suicide, and those who have serious difficulties in relationships with the opposite sex and in social situations (Maris 2002).

Death by suicide is much more difficult and tragic than death by natural causes. It leaves serious psychological effects on many people, from the family of the person to the society in which the person lives and even the treatment team (Atlı et al. 2009). Therefore, it is extremely important for clinicians to be able to recognize and prevent the risk of suicide in advance. However, reasons such as the biological, sociological and psychological components of suicidal behavior and the fact that most of those who commit suicide do not have a previous mental illness or treatment history cause limitations in preventing suicide. On the other hand, this limitation has led researchers to identify biological markers that predict suicide.

Many studies conducted in recent years have drawn attention to the role of inflammatory processes (Al-Fadhel et al. 2019, Diem et al. 2017, Gao et al. 2015, Özdin & Böke 2019), which are involved in the pathophysiology of chronic diseases, including mental illnessses, in the etiology of suicide and argued that the biological structures involved in this study can be used to predict the risk of suicide attempt and even to determine the severity of the attempt (Jha et al. 2020, Orum et al. 2018). It has been determined that inflammatory markers such as C-reactive protein (CRP), TNF-α, IL-6 are increased in the cerebrospinal fluid and plasma of patients with suicidal behavior (O'Donovan et al. 2013, Velasco et al. 2020), Neutrophil-Lymphocyte Ratio (NLR) is associated with suicidality in major depressive patients (Ekinci & Ekinci 2017), platelet volume and the demonstration of the relationship between Platelet-Lymphocyte Ratio (PLR) and the severity of suicidal behavior also supported this hypothesis (Orum et al. 2018).

Similar support has come from studies reporting an imbalance between proinflammatory and anti-inflammatory cytokines in postmortem brain tissue of people who died by suicide, compared with controls (Pandey et al. 2018). This effect of inflammation on suicidality is explained by peripheral inflammatory cytokines reaching the central nervous system via the blood-brain barrier, where monoamine neurotransmission and neuroplasticity are impaired (Miller et al. 2013).

Despite all this information, the prediction of suicidal behavior is still made with partially reliable scales, and these biomarkers are not used in clinical practice (Quinlivan et al. 2016). However, PLR, NLR, Monocyte Lymphocyte Ratio (MLR) are easily obtainable, inexpensive and reproducible values from whole blood cell (CBC) counts, and analytical techniques with high reliability that will facilitate the practical evaluation of these data can guide healthcare professionals in the prediction and prevention of suicide (Jones et al. 1999).

Numerous recent studies have shown that machine learning (ML) algorithms can be used successfully to detect predictors of suicide (Belsher et al. 2019, Choi et al. 2018, Kessler et al. 2015, Simon et al. 2018). In health, machine learning includes a series of applications that reveal biological clues to the etiology of diseases by establishing relationships between huge data that brings together demographic and physical characteristics of patients, laboratory and imaging, genetic and EEG findings (Cabitza & Banfi 2018, Percin et al. 2019, Spijker et al. 2010, Yagin et al. 2021). On the other hand, applications of machine learning are limited to black box models with little interpretability, which is one of the most important obstacles to its adoption in practical application. For this reason, there is a need for methods that will make the outputs transparent and present the predictions with their reasons to the user and allow them to be interpreted more easily, and explainable artificial intelligence (xAI) models are capable of meeting this need (Gunning et al. 2019, Tjoa & Guan 2020). However, there is no study on suicide and xAI in the literature.

In the current study, we aimed to predict the risk and severity of suicide attempts using CBC values and to develop an explainable artificial intelligence model to identify the CBC candidate that may be associated with them. This incremental model, developed using machine learning models, will not only increase the accuracy of estimating suicide risk, but also help clinicians better understand the decision-making process for assessing suicide severity, efficiently identify individuals at risk in both the psychiatric and general population, and provide early intervention.

SUBJECTS AND METHODS

Participants and study protocol

The approval for the study was obtained from the Inonu University Health Sciences Non-interventional Clinical Research Ethics Committee, Malatya, Turkey (2021/2098). In this retrospective study, medical data of 747 patients over the age of 18 who applied to the emergency unit due to suicide in the last 10 years were analyzed. The demographic characteristics of these patients, the type of suicide attempts and the data on the hemogram values [White Blood Cell (WBC), Red Blood Cell (RBC), Hemoglobin (HGB), Basophil (BA), Eosinophil (EO), Lymphocyte (LY), Monocyte (MO), Neutrophil (NEU), Mean Platelet Volume (MPV), Platelet (PLT), Hematocrit (HCT), C-Reactive Protein (CRP)] at the time of admission were analyzed and recorded by two psychiatrists. 203 patients who received immunosuppressive therapy, had acute or chronic inflammatory diseases (rheumatoid arthritis, Crohn's, systemic lupus erythematosus, chronic liver and kidney diseases, chronic heart diseases, anemia, polycythemia vera, malignancy, pneumonia, tract infection, gastroenteritis, etc.), had a history of recent surgical intervention, pregnants and were missing data were not included in the study. The control group, on the other hand, consisted of 458 healthy individuals without any physical or psychiatric disease who applied to the board of health committees to receive a clean bill of health between similar dates and were matched with the patient group in terms of age and gender. In addition, inspired by previous studies patients who committed suicide were classified as nonviolent suicide attempt group (NVSA) (overdose, poisoning, and carbon monoxide asphyxiation) and violent suicide attempt group (VSA) (hanging, cutting, drowning, other) according to the severity of the suicide attempt classified (Conwell et al. 1990).

The sample size was determined by using the total population sampling method, taking into account the exclusion criteria between the specified dates.

Statistical analysis

IBM SPSS Statistics 26.0 software was used for analysis. Since the quantitative data did not show a normal distribution, the median (minimum-maximum) and qualitative data were summarized as frequency (percentage). Mann-Whitney U test and Chi-square tests were used where appropriate for statistical analysis. $p<0.05$ was considered statistically significant.

Proposed Methodology based on Explainable Artificial Intelligence for Suicide Prediction

Predicting suicide based on xAI using the dataset described in the previous topic is the main topic of this article. The xAI procedure for suicide consists of two main steps: (i) generating different AI models to obtain the optimal model for suicide and suicide severity estimation, and (ii) using the SHapley Additive exPlanations (SHAP) approach for clinical interpretation of the optimal model.

Data preprocessing

Artificial intelligence (AI) algorithms make weaker and more biased predictions than there is a class imbalance problem in the dataset (for example, when there are 50 samples in the patient group and 500 samples in the control group). Resampling is one of the commonly preferred approaches for dealing with an unbalanced dataset. There are generally two types of methods for this, namely undersampling and oversampling. In most cases, oversampling methods are preferred over undersampling methods. This is because undersampling tends to remove samples from the data that might carry some important information. The Synthetic Minority Over-sampling Technique (SMOTE) is an oversampling method in which synthetic samples are produced for the minority class and creates virtual training records with linear interpolation for the minority class. These synthetic training records are generated by randomly selecting one or more k-nearest neighbors for each sample in the minority class. After oversampling, the data is reconstructed and various classification models can be applied to the processed data. In the data set used in this study, SMOTE (Chawla et al. 2002) oversampling method was used to balance the observations that committed suicide (544) and non-suicidal (458) observations. As a result of SMOTE, a total of 1088 data were obtained, with 544 suicidal and 544 non-suicidal. In the group created to evaluate the severity of suicide, there were 62 patients with VSA and 482 patients with NVSA, and the data set after SMOTE was balanced, resulting in a total of 964 data, 482 in each group.

Training and evaluation of models

For the suicide dataset based on blood parameters, two different models were made: suicidal/non-suicidal and VSA/NVSA. For this purpose, the balanced data sets were randomly divided into two as 80% training and 20% test set. Random Forest (RF), Logistic Regression (LR), Support vector machines (SVM) and eXtreme Gradient Boosting (XGBoost) algorithms were used in the modeling phase. The RF algorithm can be used in both classification and regression problems. In medicine, the RF algorithm is used for purposes such as identifying the correct combination of ingredients in drugs and identifying diseases by analyzing the patient's medical records. There are two stages in the RF algorithm, the first is to generate a RF, and the second is to make predictions over the RF classifier created in the first stage (Qi 2012). SVM is a machine learning method that applies the inductive principle of structural risk minimization in order to obtain a good level of generalization over a limited number of learning patterns. It is a successful algorithm that is frequently used in classification and regression problems such as RF in SVM. In this algorithm, each data item is plotted as a point in the n-dimensional space with the value of each feature being the value of a particular coordinate. Then, classification is performed by finding the hyperplane that distinguishes the two classes quite well (Wang & Hu 2005). LR can be thought of as a regression problem where the dependent variable is a categorical variable. It is widely used in linear classification problems. Although it is called regression, there is a classification here. The purpose of the LR Analysis is to establish a biologically acceptable model that can describe the relationship between dependent and independent variables in a way that has the best fit by using the least variable (Dreiseitl & Ohno-Machado 2002).

XGBoost is a high-performance version of the Gradient Boosting algorithm optimized with various modifycations. The most important features of the algorithm are its ability to obtain high predictive power, to prevent over-learning, to manage incomplete observations and to do these very quickly. The method works 10 times faster than other popular algorithms (Chen et al. 2015, Ogunleye & Wang 2020). Hyperparameters are adjustable parameters that can allow to control the training process of ML models. Model performance in ML is highly dependent on hyperparameters. Hyperparameter optimization, on the other hand, is the process of finding the configuration of the hyperparameters of the relevant model to get the best performance from an ML model. Grid search method is a successful method that is frequently used in hyperparameter optimization of ML models. Cross validation is often used to select the optimal model in the method. Cross validation is a model validation technique that tests the predictions that a ML model will make on an independent data set. This technique increases the generalizability of the model by eliminating overfitting and selection bias in model selection. In this study, the optimal hyper-parameters of each model were determined by Grid Search method using 5 replicates and 5-fold k-fold Cross Validation. With the parameters determined for each ML method, RF, LR, SVM and XGBoost models were trained on the training set. Then, the performances of the created models were evaluated on the test set and the results were compared. F1 score, accuracy, specificity and sensitivity criteria were used to evaluate the performance of the models. After a comprehensive evaluation of these performance criteria, the best performing model among the four models was selected. Finally, clinical explanations of the optimal model were created with the SHAP method.

Explainable Artificial Intelligence and the Clinical Explanation of the Optimal Model

AI methods are often referred to as "black boxes" because understanding why or how an algorithm produces accurate predictions for a given group of patients can be clinically difficult. The complexity that comes with the success of AI applications makes it very difficult to explain. xAI that emerged to solve these problems is a conceptual name given to all of the methods or techniques that aim to make various AI applications understandable by their users (Adadi & Berrada 2018, Tosun et al. 2020). In this study, SHAP, one of the explainable artificial intelligence methods, was used to find important biomarkers of suicide and suicide attempt severity and to explain/interpret the optimal model. SHAP is a new way to describe various black-box AI models developed by Lundberg and Lee as a unified framework for interpreting AI predictions (Lundberg & Lee 2017). Compared to other methods, SHAP can provide simultaneous local (personal) and global (all data related) interpretability and has a strong theoretical foundation (Abdollahi & Pradhan 2021, Hall et al. 2019). A positive SHAP value indicates that the relevant variable has a positive contribution to the target variable (here, suicide), while a negative SHAP value indicates that the related variable's contribution is negative. These contributions are shown in decreasing order of importance of the variables.

RESULTS

The study included 544 patients and 458 control cases. The median age (min-max) of the suicide group was 30 years (18-81) and 29 years (19-74) in the control group (*p=0.229*). The suicide group consisted of 336 (53.2%) female and 208 (56.1%) male patients, while the control group consisted of 295 females (46.8%) and 163 (43.9%) males (*p=0.388*). While 290 patients who attempted suicide had a known history of mental illness, 254 patients did not.

Table 1 presents the descriptive statistics of the methods used in suicide attempts. The most common preferred method for suicide was taking high-dose drugs; 470 (86.4%). The least used method was drowning; 1 (0.2%) (Table 1).

Table 1. Methods used in suicide attempts

Suicide method	n	%
Overdose drug	470	86.4
Cutting	10	1.8
Hanging	11	2.0
Gun	11	2.0
Jumping from height	14	2.6
Pesticide	2	0.4
Rat poison	2	0.4
Ingesting Corrosive	5	0.9
Ingesting insecticide	1	0.2
Piercing tool	6	1.1
Self-burning	9	1.7
Drowning	1	0.2
Gas poisoning	2	0.4

The results of the comparison of the CBC values of the suicide and control groups are given in Table 2. In the comparison between the CBC values of the patient and control groups, WBC, EO, LY, MO, NEU, NLR, MLR levels were higher in the patient group than in the control group (*p<0.05*). In the healthy group, RBC, HGB, PLR, and HCT levels were higher than those in the suicide group (*p<0.05*) (Table 2).

Table 2. Comparison of the CBC values of the suicide and control groups

Variables	Groups - Median (Min-Max) Suicide	Groups - Median (Min-Max) Control	*p*	effect size
WBC	9.4 (4.4-42)	7.55 (3.51-15.83)	*<0.001*	*2.865*
RBC	4.83 (2.9-8.42)	4.97 (3.9-7.22)	*<0.001*	*0.304*
HGB	13.9 (4.5-19.6)	14 (6.6-18.1)	*0.007*	*0.171*
BA	0.04 (0-0.28)	0.04 (0-0.6)	0.387	-
EO	0.1 (0-4.82)	0.11 (0-2.02)	*<0.001*	*0.228*
LY	2.47 (0.42-10.24)	2.33 (0.68-5.33)	*0.007*	*0.170*
MO	0.7 (0.11-2.83)	0.56 (0.23-1.27)	*<0.001*	*0.704*
NEU	6.42 (0-38.42)	4.315 (1.57-11.29)	*<0.001*	*1.096*
MPV	10.1 (0.57-87.4)	10.2 (0-13.35)	0.176	-
PLT	267 (31-748)	266 (10.9-555)	0.126	-
NLR	2.53 (0-24.57)	1.86 (0.52-9.36)	*<0.001*	*0.663*
MLR	0.271 (0.05-2.26)	0.238 (0.09-0.87)	*<0.001*	*0.390*
PLR	106.97 (8.01-889.47)	115.63 (3.33-350)	*0.001*	*0.208*
HCT	41.2 (4.73-70.5)	42.3 (27.1-54.1)	*<0.001*	*0.314*

Min-Max: Minimum-Maximum; CBC: complete blood count; WBC: Leukocyte (g/dl); RBC: erythrocyte; HGB: hemoglobin (g/dl); BA: basophil; EO: eosinophil; LY: lymphocyte; MO: monocyte; NEU: neutrophil; MPV: mean platelet volume; PLT: platelet; NLR: Neutrophil/Lymphocyte ratio; MLR: Monocyte/Lymphocyte ratio; PLR: Platelet/Lymphocyte ratio; HCT: hematocrit

Table 3. The relationship between violence of suicide attempt and demographic characteristics

	Suicide method				p	effect size
	NVSA Mean ± SD		VSA Mean ± SD			
Age	32.88±11.6		37.56±15.5		*0.004***	0.425
	n	%	n	%		
Gender					*<0.001***	*0.265*
Female	320	66.4	16	25.8		
Male	162	33.6	46	74.2		
Marital status					0.341*	-
Married	144	55.8	13	41.9		
Single	108	41.9	17	54.8		
Divorced/widowed	6	2.3	1	3.2		
Insurance					0.543*	-
No	3	0.6	1	1.6		
Yeşil kart[#]	114	23.7	17	27.4		
Social Insurance Institution	365	75.7	44	71		
History of alcohol/substance abuse					*0.049***	*0.092*
No	403	83.6	45	72.6		
Yes	79	16.4	17	27.4		
History of psychiatric disorder					0.286*	-
No	229	47.5	25	40.3		
Yes	253	52.5	37	59.7		
Preliminary-diagnosis					*0.026***	*0.266*
Depression	19[a]	10.8	3[a]	11.1		
Diagnosis was delayed	113[a]	64.2	13[a]	48.1		
Anxiety disorders	17[a, b]	9.7	4[a, b]	14.8		
Psychotic disorders	5[b]	2.8	5[b]	18.5		
Personality disorders/ addiction	15[a]	8.5	2[a]	7.4		
Bipolar disorder	3[a, b]	1.7	0[a, b]	0.0		
Mental retardation	4[a, b]	2.3	0[a, b]	0.0		

[a, b] Different characters in each row show a statistically significant difference ($p<0.05$); SD: Standard deviation;
* Pearson Chi-square test; ** Independent Samples t Test; NVSA: Nonviolent Suicide Attempt; VSA: Violent Suicide Attempt; Yeşil kart[#]: health card for uninsured people in Turkey

Table 4. Comparison of violence of suicide attempt and CBC values

Variables	Method of suicide attempt - Median (Min-Max) NVSA	VSA	*p*	effect size
WBC	9.37 (4.4-36.32)	9.885 (6.57-42)	*0.009*	*0.224*
RBC	4.82 (3.59-6.64)	4.95 (2.9-8.42)	0.652	-
HGB	13.9 (4.5-18.9)	14.2 (9.2-19.6)	0.884	-
BA	0.04 (0-0.21)	0.04 (0.01-0.28)	0.263	-
EO	0.095 (0-3.88)	0.1 (0-4.82)	0.281	-
LY	2.445 (0.42-10.24)	2.905 (0.67-8.36)	0.183	-
MO	0.68 (0.11-2.26)	0.805 (0.28-2.83)	*0.001*	*0.296*
NEU	6.365 (0-23.87)	7.96 (2.4-38.42)	*0.002*	*0.266*
MPV	10.1 (0.57-87.4)	10.28 (8.5-14.1)	0.164	-
PLT	269 (34-748)	235.5 (31-576)	*0.024*	*0.194*
NLR	2.494 (0-24.57)	3.15 (0.414-20.442)	0.124	-
MLR	0.269 (0.05-2.262)	0.32 (0.076-1.373)	*0.022*	*0.197*
PLR	108.297 (21.779-889.474)	98.679 (8.014-432.836)	*0.049*	*0.169*
HCT	40.9 (4.73-55.4)	42.15 (26-70.5)	0.114	-
CRP	0.3 (0.01-36.4)	0.33 (0.01-36.1)	*0.001*	*0.318*

Min-Max: Minimum-Maximum; NVSA: Nonviolent Suicide Attempt; VSA: Violent Suicide Attempt; CBC: complete blood count; WBC: Leukocyte (g/dl); RBC: erythrocyte; HGB: hemoglobin (g/dl); BA: basophil; EO: eosinophil; LY: lymphocyte; MO: monocyte; NEU: neutrophil; MPV: mean platelet volume; PLT: platelet; NLR: Neutrophil/ Lymphocyte ratio; MLR: Monocyte/Lymphocyte ratio; PLR: Platelet/Lymphocyte ratio; HCT: hematocrit; CRP: C-reactive protein

The relationship between demographic variables and violence of suicide attempt is shown in Table 3. At the initial psychiatric evaluation, of the patients who attempted non-violent suicide, 113 did not have a specified diagnosis. 5 patients were diagnosed with psychosis and 17 patients with anxiety disorder. Being male (*p<0.001*) and older age (*p=0.004*) increased the rates of VSA. The rate of attempting violent suicide was higher in those with a history of alcohol-substance addiction than in those without alcohol-substance addiction (*p=0.049*) (Table 3).

Comparison between violence of suicide attempt and CBC values are given in Table 4. In the comparison between the CBC values of the VSA and NVSA groups: while WBC, MO, NEU, MLR, CRP levels of the VSA group were higher than the NVSA group (*p<0.05*), PLT and PLR levels were lower (*p<0.05*) (Table 4).

The estimation results of the models for the suicide and control groups are summarized in the confusion matrix in Table 5 (a), and the prediction results of the models for the VSA and NVSA groups are summarized in the confusion matrix in Table 5 (b). When the confusion matrices of the suicide and control groups in Figure 1(a) are examined, it was seen that the LR model correctly classified 171 of 218 patients, the RF model 165 of 218 patients, the SVM model 169 of 218 patients, and the XGBoost model correctly classified 181 of 218 patients. When the confusion matrices are examined for the estimation of the severity of the suicide attempt, the LR model was able to correctly classify 127 of 193 patients, the RF model 180 of 193 patients, the SVM model 181 of 193 patients, and the XGBoost model 182 of 193 patients (Table 5).

In Table 6; Accuracy, specificity, sensitivity and F1-score values of four models developed for suicide/control groups and violent/non-violent suicide attempt groups are given. When the performance criteria of the models created for suicide prediction are examined, XGBoost was the best performing model with 0.83 (0.78-0.88) Accuracy, 0.804 (0.752-0.857) F1-Score, 0.776 (0.68-0.854) Sensitivity and 0.875 (0.802-0.928) Specificity values. When the performance criteria of

Table 5. Confusion matrices of established models

Prediction	Reference		
Confusion matrices of established models for the suicide and control groups			
	Suicide	Control	Sum
LR			
Suicide	71	20	91
Control	27	100	127
Sum	98	120	218
RF			
Suicide	72	19	91
Control	34	93	127
Sum	106	112	218
SVM			
Suicide	71	20	91
Control	29	98	127
Sum	100	118	218
XGBoost			
Suicide	76	15	91
Control	22	105	127
Sum	98	120	218
Confusion matrices of established models for violent and nonviolent suicide attempting groups			
	VSA	NVSA	Sum
LR			
VSA	69	22	91
NVSA	44	58	102
Sum	113	80	193
RF			
VSA	82	9	91
NVSA	4	98	102
Sum	86	107	193
SVM			
VSA	88	3	91
NVSA	9	93	102
Sum	97	96	193
XGBoost			
VSA	80	9	89
NVSA	2	102	104
Sum	82	111	193

NVSA: Nonviolent Suicide Attempt; VSA: Violent Suicide Attempt

Table 6. Performance criteria of models created for suicide/control groups and violent/nonviolent suicide attempting groups

Metric	LR	RF	SVM	XGBoost
Suicide and Control Groups				
Accuracy	0.784 (0.73-0.839)	0.757 (0.7-0.814)	0.775 (0.72-0.831)	0.83 (0.78-0.88)
F1-Score	0.751 (0.694-0.809)	0.731 (0.672-0.79)	0.743 (0.685-0.801)	0.804 (0.752-0.857)
Sensitivity	0.724 (0.625-0.81)	0.679 (0.582-0.767)	0.71 (0.611-0.796)	0.776 (0.68-0.854)
Specificity	0.833 (0.754-0.895)	0.83 (0.748-0.895)	0.831 (0.75-0.893)	0.875 (0.802-0.928)
Violent and Nonviolent Suicide Attempting Groups				
Accuracy	0.658 (0.591-0.725)	0.933 (0.897-0.968)	0.938 (0.904-0.972)	0.943 (0.91-0.976)
F1-Score	0.676 (0.61-0.742)	0.927 (0.89-0.963)	0.936 (0.902-0.971)	0.936 (0.901-0.97)
Sensitivity	0.611 (0.514-0.701)	0.953 (0.885-0.987)	0.907 (0.831-0.957)	0.976 (0.915-0.997)
Specificity	0.725 (0.614-0.819)	0.916 (0.846-0.961)	0.969 (0.911-0.994)	0.919 (0.852-0.962)

LR: Logistic Regression; RF: Random Forest; SVM: Support vector machines; XGBoost: eXtreme Gradient Boosting

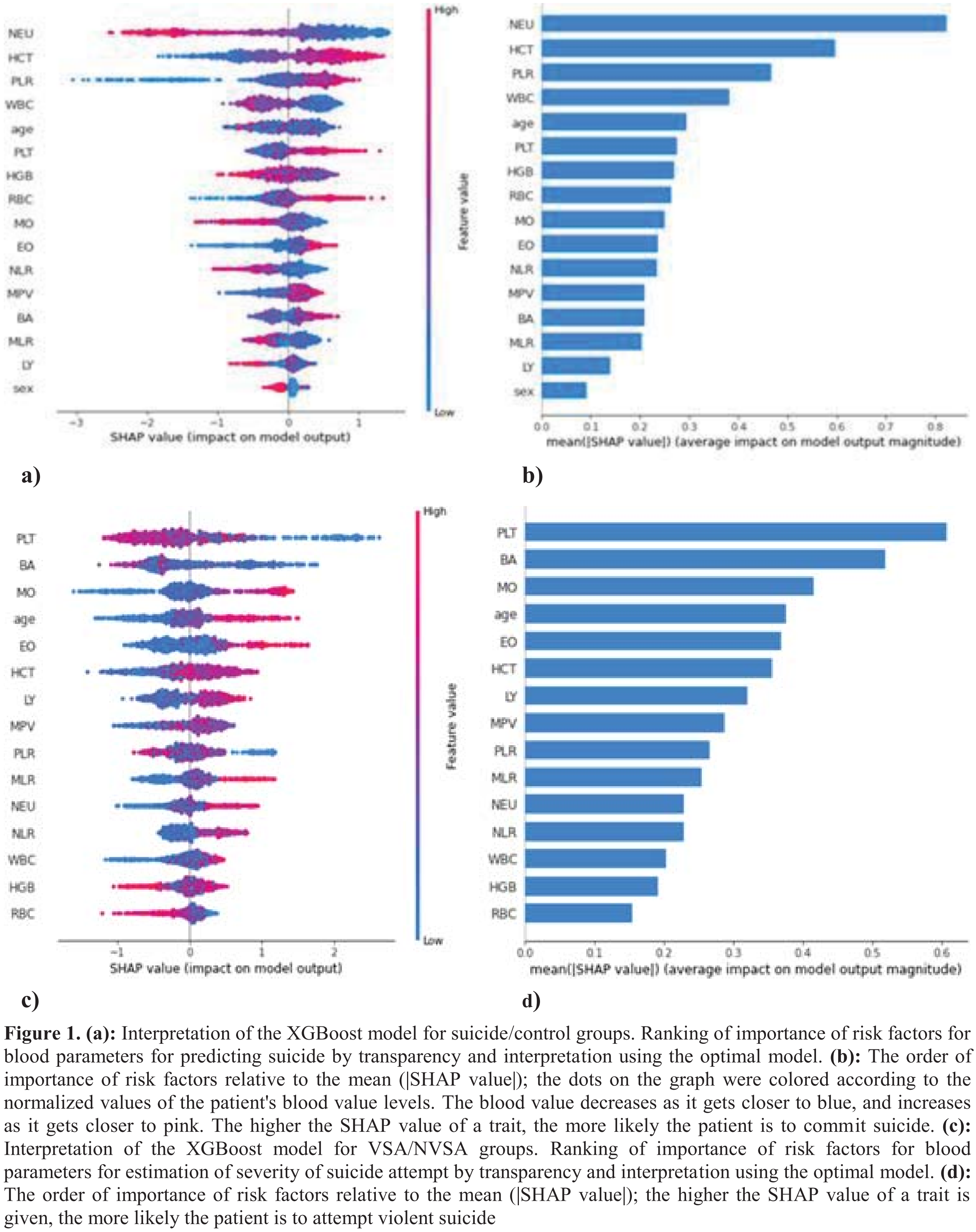

Figure 1. (a): Interpretation of the XGBoost model for suicide/control groups. Ranking of importance of risk factors for blood parameters for predicting suicide by transparency and interpretation using the optimal model. **(b):** The order of importance of risk factors relative to the mean (|SHAP value|); the dots on the graph were colored according to the normalized values of the patient's blood value levels. The blood value decreases as it gets closer to blue, and increases as it gets closer to pink. The higher the SHAP value of a trait, the more likely the patient is to commit suicide. **(c):** Interpretation of the XGBoost model for VSA/NVSA groups. Ranking of importance of risk factors for blood parameters for estimation of severity of suicide attempt by transparency and interpretation using the optimal model. **(d):** The order of importance of risk factors relative to the mean (|SHAP value|); the higher the SHAP value of a trait is given, the more likely the patient is to attempt violent suicide

the models created for suicide severity are examined, XGBoost has been the model with the best performance with 0.943 (0.91-0.976) Accuracy, 0.936 (0.901-0.97) F1-Score, 0.976 (0.915-0.997) Sensitivity and 0.919 (0.852-0.962) Specificity values (Table 6).

Figures 1 (a) and (b) show the positive or negative contributions of biomarker candidate values for suicide to the prediction of the optimal model (XGBoost). When Figure 1 (b) is examined; NEU, HCT, PLR, WBC, age, and PLT were found to be the six most important blood parameters contributing to determine the suicide risk. In Figure 1 (a), it can be said that lower NEU, WBC, age, MO, NLR and MLR levels as well as higher HCT, PLR, PLT, HGB, RBC, EO, MPV and BA levels have positive contributions to the prediction performance of the XGBoost model established for suicide risk. When Figure 1 (c) and (d) are examined, it is seen that PLT, BA, MO are the three most important blood parameters contributing to the target variable (violence of suicide attempt). In addition, lower levels of PLT, BA, PLR, and RBC, and higher levels of MO, age, EO, HCT, LY, MLR, NEU, NLR, WBC, and HGB have positive contributions to the prediction of the patient's model to predict violent suicide (Figure 1).

DISCUSSION

This study has two innovative sides. First, the current study is the first to use ML based on CBC values to distinguish both the risk of suicide and the severity of suicidal behavior. Second, CBC values that predict suicide risk and severity of suicidal behavior are presented for the first time with a methodology combining ML and xAI. The results show that this approach combining ML and xAI can assist clinicians in identifying individuals at risk.

In the first phase of our study, a large suicide population was compared with a sex-age-matched control group using 4 different ML methods. The primary aim was to detect suicide with the most accurate outcome. For this purpose, in addition to RF, LR, SVM, which are well-known ML methods, the XGBoost method, which gives very successful results, was also used. The grid search method for hyper parameter optimization was used to improve the performance of ML models. Among the four ML models, the model that performed the best in the test set was selected to investigate the explainability of the model, and SHAP which was the xAI method was used to interpret the outputs of this model. In this way, in addition to obtaining the importance of CBC values, it was also possible to determine which levels of these values (low, high) were more closely related to suicide. This is an important result and may help clinicians determine which CBC values should be considered first in patients who come to the clinic with a suicide attempt. In the second stage of the study, the cases that committed suicide were divided into two groups as VSA and NVSA. Methods in the first stage were used to determine the severity of the suicide attempt.

Although there are many studies showings that hemogram parameters can be used to determine suicide risk, the number of studies this evaluation with ML is very limited. Arı et al., with the method they developed in their study, estimated the suicide risk in adolescents with an accuracy of 93.5% using hemogram parameters (Arı et al. 2020). However, only the performance of the ML model was examined in the study, and biomarkers that could be effective for suicide were not examined. In the current study, in addition to the estimation of suicide risk and severity of adult individuals, biomarkers based on hemogram were investigated.

As a result of the study, we found that XGBoost model was the best performing model to predict suicide risk and suicide attempt severity among the four ML methods applied to CBC data. While the XGBoost model distinguished suicidal individuals from healthy individuals with an accuracy of 0.83 (0.78-0.88), it was able to distinguish between violent and nonviolent groups with an accuracy of 0.943 (0.91-0.976).

The relationship between suicide attempt and inflammatory response was also investigated in this study; In patients who attempted suicide, the inflammatory response was increased and the levels of WBC, EO, LY, MO, NEU and MLR, including NLR, were found to be high. These results supported studies that argued that circulating inflammatory marker levels were higher in suicidal individuals as an indicator of the activation of the immune system and that suicide should now be considered as a pro-inflammatory state (Black & Miller 2015, Russell et al. 2021). As a matter of fact, in these studies, it was shown that neutrophils, monocytes, cells that initiate acute inflammation, and lymphocyte activation, which are protective regulators of the immune system, are associated with a series of psychiatric diseases, including suicide, by causing activation of pro inflammatory cytokines and increased serum levels (which affect the related areas negatively by passing on to the CNS) (Puangsri & Ninla-Aesong 2021). In addition, the researchers stated that all these parameters can be used to predict and manage individuals at risk. Ekinci et al. showed that NLR may be a marker for suicidality in patients with major depression (Ekinci & Ekinci 2017). Ucuz et al. found that almost all hemogram parameters showing inflammation in patients who committed suicide were significantly higher than those who did not commit suicide (Ucuz &Tetik 2020). Another recent study Keaton SA et al. also found that increased WBC values were associated with suicide in patients with depression (Keaton et al. 2019). While Endres et al. showed that suicidality (thoughts and behavior) in depressed patients was associated with increased white blood cell count (Endres et al. 2016). Batty et al. reported that higher levels of white blood cell counts were a prelude to later suicide (Batty et al. 2018). The researchers further demonstrated that there is a bidirectional relationship between suicide and inflammation, that suicidal thoughts and its provoking factors activate biological systems, including the inflammatory response, and/or inflammation by creating a behavioral pattern that includes symptoms of depression such as a lack of pleasure, hopelessness, and fatigue etc. that can lead to illness depression and trigger suicidal thoughts (O'Donovan et al. 2013). In the present study, however, since diseases that may affect the inflammatory response were excluded, the results obtained provided new evidence for the view that suicide attempt increases inflammation.

An important aspect of this study was that it presented a model that evaluated a large number of variables that could be used in the estimation of suicide risk and severity, and their predictive power. While current studies investigating CBC biomarkers of suicide try to determine only the high or low level of CBC for suicide, thanks to the model we created in this study, we had the chance to list the most important CBC values. In addition, we found that high or low levels of the most significant CBC values together may more accurately predict suicide. Accordingly, we proved that low NEU, WBC, MO, NLR, MLR levels and high HCT, PLR, PLT, HGB, RBC, EO, MPV, BA levels and being young increase the accuracy of suicidal risk estimation.

On the other hand, our results could not fully support studies showing a relationship between the violence of suicide attempt and higher NEU, MPV, NLR, PLR, and lower LY (Aguglia et al. 2021, Orum et al. 2018). In our study, WBC, MO, NEU, MLR, CRP levels were higher in those who showed violent suicidal attempted than those in the nonviolent suicidal behavior group. Contrary to expectations, PLT and PLR levels were lower and NLR and LY levels did not have any effect. The CRP level of the VSA group was high, and this result supported the study of Kumar et al., which showed that CRP could be used to detect suicide in patients with depression (Kumar et al. 2021). The developed model determined the three most important values for the severity of the suicide attempt as PLT, BA, and MO. In addition, lower values of PLT, BA, PLR and RBC levels and higher values of MO, EO, HCT, LY, MLR, NEU, NLR, WBC, HGB levels and older age of the patient have positive contributions to the performance of the model created to predict violent suicide. It was thought that this difference between studies might be related to many factors, including the time of blood collection, the method of collection, the laboratory method applied, and the types of suicide attempts. For example, in our study, the blood of the patients was taken within the first hour of their admission to the emergency room, whereas in the Orum et al study (Orum et al. 2018), the blood was taken after 8 hours of fasting and in the morning hours, and this may have been the reason for the difference. Nevertheless, this result was important in demonstrating the need for studies to be conducted in larger patient populations and within predetermined standards in order to develop a firm judgment.

This study has some methodological limitations. First, the data were obtained from electronic health records, the quality of which may limit results. Second, the retrospective nature of the study precludes establishing a definitive causality between suicide and hemogram values. Therefore, prospective studies are required to confirm causality. Third, parameters that could affect hemogram values such as body mass index (BMI), type of medicine could not be evaluated and patients with a diagnosis of psychiatric illness were not excluded, as these could be associated with inflammatory processes. Despite all these limitations, this study is the first study based on xAI that evaluates hemogram values for suicide and suicide attempt severity and will serve as a basic reference in the country. An age- and sex-matched control group from the healthy population strengthened our assessment. In general, the sample size was sufficient and consisted of data selected from both tertiary and secondary health centers, which increased the generalizability of the results of this study.

CONCLUSION

As a result, CBC values can be used as a biomarker to determine suicide risk and severity. XGBoost from ML methods obtains strong predictions for suicide. ML and xAI, their applications may be useful in future therapeutic targets in suicide detection and prevention and in personalized medicine choices.

Acknowledgements: None.

Conflict of interest: None to declare.

Contribution of individual authors:

Neslihan Cansel: concept and design of the article; literature searches; writing and manuscript.

Fatma Hilal Yagin: performed the statistical analysis; training and evaluation of models.

Mustafa Akan & Bedriye Ilkay Aygul: collected the data; manuscript preparation.

All authors have read and approved the manuscript.

References

1. *Abdollahi A & Pradhan B: Urban vegetation mapping from aerial imagery using Explainable AI (XAI). Sensors 2021; 21:4738. https://doi.org/10.3390/s21144738*
2. *Adadi A & Berrada M: Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE access 2018; 6:52138-52160*
3. *Aguglia A, Amerio A, Asaro P, Caprino M, Conigliaro C, Giacomini G et al.: High-lethality of suicide attempts associated with platelet to lymphocyte ratio and mean platelet volume in psychiatric inpatient setting. World J Biol Psychiatry 2021; 22:119-127*
4. *Al-Fadhel SZ, Al-Hakeim HK, Al-Dujaili AH & Maes M: IL-10 is associated with increased mu-opioid receptor levels in major depressive disorder. Eur Psychiatry 2019; 57:46-51*
5. *Arı B, Arı A & Şengür A: Suicide Prediction from Hemogram with Machine Learning. Avrupa Bilim ve Teknoloji Dergisi 2020; 364-369*

6. Atlı Z, Eskin M & Dereboy Ç: Validity and reliability of the Suicide Probability Scale (PSS) in clinical samples. Klinik Psikiyatri Dergisi 2009; 12:111-124
7. Batty GD, Jung KJ, Lee S, Back JH & Jee SH: Systemic inflammation and suicide risk: cohort study of 419 527 Korean men and women. J Epidemiol Community Health 2018; 72:572-574
8. Belsher BE, Smolenski DJ, Pruitt LD, Bush NE, Beech EH, Workman DE et al.: Prediction models for suicide attempts and deaths: a systematic review and simulation. JAMA psychiatry 2019; 76:642-651
9. Bhatia MS, Aggarwal NK & Aggarwal B: Psychosocial profile of suicide ideators, attempters and completers in India. Int J Soc Psychiatry 2000; 46:155-163
10. Black C & Miller BJ: Meta-Analysis of Cytokines and Chemokines in Suicidality: Distinguishing Suicidal Versus Nonsuicidal Patients. Biol Psychiatry 2015; 78:28-37
11. Cabitza F & Banfi G: Machine learning in laboratory medicine: waiting for the flood?. Clin Chem Lab Med 2018; 56:516-524
12. Chawla NV, Bowyer KW, Hall LO & Kegelmeyer WP: SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002; 16: 321-357
13. Chen T, He T,Benesty M, Khotilovich V, Tang Y, Cho H & Chen K: Xgboost: extreme gradient boosting. R package version 04-2 2015; 1:1-4
14. Choi SB, Lee W, Yoon J-H, Won JU & Kim DW: Ten-year prediction of suicide death using Cox regression and machine learning in a nationwide retrospective cohort study in South Korea. J Affect Disord 2018; 231: 8-14
15. Conwell Y, Rotenberg M & Caine ED: Completed suicide at age 50 and over. J Am Geriatr Soc 1990; 38:640-644
16. Diem S, Schmid S, Krapf M, Flatz L, Born D, Jochum W et al.: Neutrophil-to-Lymphocyte ratio (NLR) and Platelet-to-Lymphocyte ratio (PLR) as prognostic markers in patients with non-small cell lung cancer (NSCLC) treated with nivolumab. Lung cancer 2017; 111:176-181
17. Dreiseitl S & Ohno-Machado L: Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 2002; 35:352-359
18. 19. Ekinci O & Ekinci A: The connections among suicidal behavior, lipid profile and low-grade inflammation in patients with major depressive disorder: a specific relationship with the neutrophil-to-lymphocyte ratio. Nord J Psychiatry 2017; 71:574-580
19. Endres D, Perlov E, Dersch R, Baumgartner A, Hottenrott T, Berger B et al.: Evidence of cerebrospinal fluid abnormalities in patients with depressive syndromes. J Affect Disord 2016; 198:178-184
20. Gao SQ, Huang LD, Dai RJ, Chen DD, Hu WJ & Shan YF: Neutrophil-lymphocyte ratio: a controversial marker in predicting Crohn's disease severity. Int J Clin Exp Pathol 2015; 8:14779-14785
21. Gunning D, Stefik M, Choi J, Miller T, Stumpf S & Yang GZ: XAI - Explainable artificial intelligence. Sci Robot 2019; 4:eaay7120. https://www.science.org/doi/10.1126/scirobotics.aay7120
22. Hall M, Harborne D, Tomsett R, Galetic V, Quintana-Amate S, Nottle A & Preece A: A systematic method to understand requirements for explainable AI (XAI) systems. In Proceedings of the IJCAI Workshop on eXplainable Artificial Intelligence (XAI 2019), Macau, China 2019
23. Jha MK, Cai L, Minhajuddin A, Fatt CC, Furman JL, Gadad BS et al.: Dysfunctional adaptive immune response in adolescents and young adults with suicide behavior. Psychoneuroendocrinology 2020; 111:104487. https://doi.org/10.1016/j.psyneuen.2019.104487
24. Jones SA, Novick D, Horiuchi S, Yamamoto N, Szalai AJ & Fuller GM: C-reactive protein: a physiological activator of interleukin 6 receptor shedding. J Exp Med 1999; 189:599-604
25. Keaton SA, Madaj ZB, Heilman P, Smart L, Grit J, Gibbons R et al.: An inflammatory profile linked to increased suicide risk. J Affect Disord 2019; 247:57-65
26. Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ et al.: Predicting suicides after psychiatric hospitalization in US Army soldiers: the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). JAMA Psychiatry 2015; 72:49-57
27. Kumar K, Srivastava S, Sharma B, Avasthi RK & Kotru M: Comparison Between Inflammatory Biomarkers (High-Sensitivity C - reactive protein and Neutrophil-Lymphocyte Ratio) and Psychological Morbidity in Suicide Attempt Survivors Brought to Medicine Emergency. Cureus 2021; 13:e17459. doi:10.7759/cureus.17459
28. Lundberg SM & Lee SI: A unified approach to interpreting model predictions. Proceedings of the 31st international conference on neural information processing systems 2017; 4768-4777
29. Maris RW: Suicide. Lancet 2002; 360:319-26
30. Miller AH, Haroon E, Raison CL & Felger JC: Cytokine targets in the brain: impact on neurotransmitters and neurocircuits. Depress Anxiety 2013; 30:297-306
31. O'Donovan A, Rush G, Hoatam G, Hughes BM, McCrohan A, Kelleher C et al.: Suicidal ideation is associated with elevated inflammation in patients with major depressive disorder. Depress Anxiety 2013; 30:307-314
32. Ogunleye A & Wang QG: XGBoost model for chronic kidney disease diagnosis. IEEE/ACM Trans Comput Biol Bioinform 2020; 17:2131-2140
33. Orum MH, Kara MZ & Egilmez OB: Mean platelet volume and neutrophil to lymphocyte ratio as parameters to indicate the severity of suicide attempt. J Immunoassay Immunochem 2018; 39:647-659
34. Özdin S & Böke Ö: Neutrophil/lymphocyte, platelet/ lymphocyte and monocyte/lymphocyte ratios in different stages of schizophrenia. Psychiatry Res 2019; 271:131-135
35. Pandey GN, Rizavi HS, Zhang H, Bhaumik R & Ren X: Abnormal protein and mRNA expression of inflammatory cytokines in the prefrontal cortex of depressed individuals who died by suicide. J Psychiatry Neurosci 2018; 43:376-385
36. Percin I, Yagin FH, Guldogan E & Yologlu S: ARM: An Interactive Web Software for Association Rules Mining and an Application in Medicine. 2019 International Artificial Intelligence and Data Processing Symposium (IDAP). IEEE 2019; 1-5
37. Puangsri P & Ninla-Aesong P: Potential usefulness of complete blood count parameters and inflammatory ratios as simple biomarkers of depression and suicide risk in drug-naive, adolescents with major depressive disorder. Psychiatry Res 2021; 305:114216. https://doi.org/10.1016/j.psychres.2021.114216

38. *Qi Y: Random forest for bioinformatics. Ensemble machine learning: Springer, Boston, MA 2012; 307-323*
39. *Quinlivan L, Cooper J, Davies L, Hawton K, Gunnell D & Kapur N: Which are the most useful scales for predicting repeat self-harm? A systematic review evaluating risk scales using measures of diagnostic accuracy. BMJ open 2016; 6:e009297. http://dx.doi.org/10.1136/bmjopen-2015-009297*
40. *Russell AE, Mars B, Wen CP, Chang SS & Gunnell D: Evidence for an association between inflammatory markers and suicide: a cohort study based on 359,849 to 462,747 Taiwanese adults. J Affect Disord 2021; 281:967-971*
41. *Simon GE, Johnson E, Lawrence JM, Rossom RC, Ahmedani B, Lynch FL et al.: Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. Am J Psychiatry 2018; 175:951-960*
42. *Spijker S, Van Zanten JS, De Jong S, Penninx BW, van Dyck R, Zitman FG et al.: Stimulated gene expression profiles as a blood marker of major depressive disorder. Biol Psychiatry 2010; 68:179-186*
43. *Tjoa E & Guan C: A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. IEEE Trans Neural Netw Learn Syst 2020; 32:4793-4813*
44. *Tosun AB, Pullara F, Becich MJ, Taylor D, Fine JL & Chennubhotla SC: Explainable AI (xAI) for anatomic pathology. Adv Anat Pathol 2020; 27:241-250*
45. *Ucuz İ & Tetik BK: Can suicide behavior and seasonality of suicide be predicted from inflammatory parameters in adolescents? Med Hypotheses 2020; 143:110061. https://doi.org/10.1016/j.mehy.2020.110061*
46. *Velasco Á, Rodríguez-Revuelta J, Olié E, Abad I, Fernández-Peláez A, Cazals A et al.: Neutrophil-to-lymphocyte ratio: a potential new peripheral biomarker of suicidal behavior. Eur Psychiatry 2020; 63:e14. https://doi.org/10.1192/j.eurpsy.2019.20*
47. *Wang H & Hu D: Comparison of SVM and LS-SVM for regression. In 2005 International Conference on Neural Networks and Brain. IEEE 2005; 1: 279-283*
48. *WHO: Suicide. Available from: https://www.who.int/newsroom/fact-sheets/detail/suicide. who. int (accessed Dec 11, 2021)*
49. *Yagin FH, Guldogan E, Ucuzal H & Colak C: A Computer-Assisted Diagnosis Tool for Classifying COVID-19 based on Chest X-Ray Images. Konuralp Medical Journal 2021; 13:438-445*

Correspondence:
Research Assistant Fatma Hilal Yagin, MD
Inonu University Faculty of Medicine, Department of Biostatistics and Medical Informatics
Malatya, Turkey
E-mail: hilal.yagin@inonu.edu.tr