ORIGINAL ARTICLE

Value of fecal calprotectin in prediction of acute appendicitis based on a proposed model of machine learning

Zeynep Küçükakçali, M.D.,[1] Sami Akbulut, M.D.,[2] Cemil Çolak, M.D.,[1]

[1]Department of Biostatistics and Medical Informatics, Inonu University Faculty of Medicine, Malatya-*Türkiye*

[2]Department of Surgery, Inonu University Faculty of Medicine, Malatya-*Türkiye*

ABSTRACT

BACKGROUND: The aim of this study is to apply random forest (RF), one of the machine learning (ML) algorithms, to a dataset consisting of patients with a presumed diagnosis of acute appendicitis (AAp) and to reveal the most important factors associated with the diagnosis of AAp based on the variable importance.

METHODS: An open-access dataset comparing two patient groups with (n=40) and without (n=44) AAp to predict biomarkers for AAp was used for this case−control study. RF was used for modeling the data set. The data were divided into two training and test dataset (80:20). Accuracy, balanced accuracy (BC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) performance metrics were appraised for model performance.

RESULTS: Accuracy, BC, sensitivity, specificity, PPV, NPV, and F1 scores pertaining to the RF model were 93.8%, 93.8%, 87.5%, 100%, 100%, 88.9%, and 93.3%, respectively. Following the variable importance values regarding the model, the variables most associated with the diagnosis and prediction of AAp were fecal calprotectin (100 %), radiological imaging (89.9%), white blood test (51.8%), C-reactive protein (47.1%), from symptoms onset to the hospital visit (19.3%), patients age (18.4%), alanine aminotransferase levels >40 (<1%), fever (<1%), and nausea/vomiting (<1%), respectively.

CONCLUSION: A prediction model was developed for AAp with the ML method in this study. Thanks to this model, biomarkers that predict AAp with high accuracy were determined. Thus, the decision-making process of clinicians for diagnosing AAp will be facilitated, and the risks of perforation and unnecessary operations will be minimized thanks to the timely diagnosis with high accuracy.

Keywords: Acute appendicitis; fecal calprotectin; machine learning; random forest; variable importance.

INTRODUCTION

Acute appendicitis (AAp) is one of the leading causes of admission to emergency departments for abdominal pain, and appendectomy is one of the most commonly performed emergency procedures globally.[1,2] Epidemiological studies have shown that the incidence of AAp varies between 100 and 233/1,00,000 population.[3-5] It was reported that the lifetime risk of AAp was around 8.6% and 6.7% in men and women, respectively.[6-8] Although it was claimed that conservative treatment consisting of antibiotic combination and close follow-up was advantageous in the management of patients with AAp, appendectomy remains the standard gold treatment in the treatment of AAp.[9,10]

The diagnosis of AAp is confirmed by the combined evaluation of anamnesis, clinical symptoms, physical examination findings, biochemical blood parameters, and radiological examinations.[11] However, despite the use of many clinical and radiological diagnostic modalities, some of the patients may

Cite this article as: Küçükakçali Z, Akbulut S, Çolak C. Value of fecal calprotectin in prediction of acute appendicitis based on a proposed model of machine learning. Ulus Travma Acil Cerrahi Derg 2023;29:655-662.

Address for correspondence: Sami Akbulut, M.D.
Inonu University Faculty of Medicine, 44280, Malatya, Türkiye
E-mail: akbulutsami@gmail.com

Ulus Travma Acil Cerrahi Derg 2023;29(6):655-662 DOI: 10.14744/tjtes.2023.10001 Submitted: 28.11.2022 Revised: 15.12.2022 Accepted: 20.02.2023

experience a diagnostic dilemma. Diagnostic uncertainties in the management of AAp pave the way for discussing two basic problems, such as unnecessary (negative appendectomy) and delayed appendectomy. Because there is much discussion in the literature about what should be the acceptable ratios for these two main problems. Although their incidence varies from center to center, it is known that negative appendectomy and perforated AAp are directly related to morbidity, mortality, and high cost.[12] It has been shown that 3.8–15% and 5.9–23.5% of children and adults who consulted the emergency department with a preliminary diagnosis of AAp were misdiagnosed, respectively.[13-15] Therefore, the correct management of patients with pre-diagnosis of AAp is one of the uttermost important issues concerning patient safety and evaluation of health care quality.[7] Until the last few years, when artificial intelligence (AI) models started to be used in the health field, conventional statistical analysis methods were used to predict AAp or perforated AAp in almost all clinical studies related to AAp. However, the most important disadvantage of these classical statistical analysis methods is that they are not technically suitable for making inferences about new data by learning from data as in data mining and AI methods. Recently, AI-based models have become popular in medical studies due to their approach to discovering hidden patterns from datasets and their ability to detect new phenomena faster and with higher accuracy (ACC) rates.[16-18]

Machine learning (ML) is a subfield of AI that employs data-driven learning to anticipate new data when exposed to new data. In recent years, ML models have been widely implemented in both illness diagnosis and systems for clinical decision support.[19] ML approaches may be used to extract valuable characteristics from retrospective structured and unstructured electronic health record data to possibly reproduce the clinician's thinking process in forecasting and detecting illnesses in emergency departments. These features have the potential to improve patient safety in emergency units by accurately identifying a patient's diagnosis.[20,21]

The random forest (RF) method used to construct a prediction set in decision trees growing in subspaces of randomly selected data, and based on these trees is a ML method.[22] RF, popularly with very successful and fast results, is often preferred in classification and regression processes.[23] The RF method can be used as an effective decision tree method in both categorical and continuous data sets, as well as large or small-sized data sets.[24] This algorithm has been demonstrated as a suitable classification method in many studies and has been stated to be an effective predictor in determining the cause and effect relationship.[23-25] This study aims to apply RF, one of the ML algorithms, to a dataset of patients with the presumed diagnosis of AAp and to reveal the factors associated with the accurate diagnosis AAp based on the variable importance obtained as a result of the model. Thus, patients can be accurately classified as with and without AAp.

MATERIALS AND METHODS

Study Design and Dataset

The current study's data were collected prospectively and examined retrospectively for modeling purposes. The study used an open-access dataset consisting of 84 patients to classify AAp and identify the related factors. The dataset is available at https://data.mendeley.com/datasets/msymp49s6s/1. The dataset was comprised of assenting adult patients (age ≥18 years) who arrived with abdominal pain at the emergency department of Wujiang District Fifth People's Hospital in 2018.[26] Demographic (age and gender), clinical (nausea/vomiting, onset to visit, and pregnancy), biochemical (white blood cell [WBC], C-reactive protein [CRP], alanine aminotransferase [ALT>40], fecal calprotectin), and radiological features (ultrasonography [US] and computed tomography [CT]) of the patients were recorded retrospectively. The exclusion criteria included patients with gastrointestinal diseases or those using nonsteroidal anti-inflammatory drugs or inhibitors of proton pump, patients with patients with inflammatory bowel disease and more than 72 h of stomach discomfort who are undergoing anti-inflammatory medication, and patients for whom stool samples were not obtained throughout the study.[26] The data set in this study consisted of 84 patients with a preliminary diagnosis of AAp, of which 40 (47.6%) patients were AAp and the remaining 44 (52.4%) were not. Table 1 explains the variables under question and their characteristics.

Random Forest

RF method is a classification and regression, including the voting method. It comprises several decision trees together, and the individual trees are voted to determine the winning class. The decision trees in the forest are unrelated of one another and are generated using the bootstrap approach and samples selected from the data set.[27] There are many classification trees in the RF method. Each of the input data is passed through all of these classification trees. Each element of the input data is classified with classification trees. After each input data is entered into all classification trees and voted, assignments are made to the class with the highest number of votes from the tree structures.[28] Determination of branching criteria and selecting a suitable pruning method in the RF method is a fundamental issue. The Gini index is utilized to calculate the RF classifier's branching criterion. The Gini index measures the degree of weakness of class attributes.[29] As with other prediction methods, the practitioner must specify the certain parameters for the RF approach. In constructing the tree structure, these parameters are the number of instances to be utilized at each node and the number of trees to be constructed. Especially, k user-specified trees are used to build the decision forest during the classification process.[30]

Table 1. Explanations of the variables in the dataset and their characteristics

Variables	Explanations of the variables	Variable type	Variable role
Group	0=Without AAp; 1=With AAp	Qualitative	Output
FC	Fecal calprotectin (µg/g)	Quantitative	Predictor
WBC	WBC	Quantitative	Predictor
CRP	CRP	Quantitative	Predictor
Age	-	Quantitative	Predictor
Gender	0=Male; 1=Female	Qualitative	Predictor
Pregnancy	0=No; 1=Yes	Qualitative	Predictor
Onset to visit	Onset to visit (h)	Quantitative	Predictor
Nausea/vomiting	0=No; 1=Yes	Qualitative	Predictor
Fever	0=No; 1=Yes	Qualitative	Predictor
ALT	ALT (0=≤40; 1=>40)	Qualitative	Predictor
Imaging	0=Negative; 1=Positive	Qualitative	Predictor

ALT: Alanine aminotransferase; FC: Fecal calprotectin; WBC: White blood cell; CRP: C-reactive protein; AAp: Acute appendicitis.

Machine Learning Modeling and Performance Evaluation

In the current study, RF was used in the modeling stage for the dataset of AAp in question. The data set was divided as 80:20 as a training and test dataset. Analyses were conducted using the n-fold cross-validation technique. In the n-fold cross-validation approach, the data are separated into n parts before the model is applied to each of the n parts. One of the n components is utilized for testing, while the remaining n-plus-one components are used to train the model. In this work, 5-fold cross-validation was performed for the modeling procedure. As performance assessment criteria, we employed accuracy (ACC), balanced accuracy (BC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1-score. In addition, variable importance was determined, which offers information on how much the factors assign importance to the outcome variable.

Protocol for Research and Ethical Committee Approval

This study, which utilized the open-access Gene Expression Omnibus dataset from the National Center for Biotechnology Information and involved human participants, was conducted in accordance with the ethical standards of the institutional and national research committees, the 1964 Helsinki Declaration and its later amendments, or other ethical standards. The Inonu University Institutional Review Board granted clearance for Non-Interventional Clinical Research (2022/3768).

Biostatistical Analysis

Since this study was based on open dataset data, the power of the study was calculated using the post hoc power analysis method. Power analysis (α=0.05, sample size=84, effect size =2.30, two-sided) using the Mann–Whitney U-test was performed, and the power (1-β) was calculated as 1.0 (100%).[31] The median (minimum-maximum) was employed to summarize the quantitative data, whereas qualitative data are presented as numbers and percentages. Shapiro–Wilk test for normality was used to identify whether the variables had a normal distribution. The Mann–Whitney U and Yates' correction chi-square tests were evaluated for a statistically significant difference between the output and input variables. $p<0.05$ values were considered statistically significant. IBM SPSS Statistics 26.0 for the Windows package program was employed in all analyses.

RESULTS

The current study consisted of 84 patients (age; mean ± standard deviation = 47.19±16.76 years) with a presumed diagnosis of AAp, 42 (50%) of whom were males and 42 (50%) were females. These patients were divided into two groups as with (n=40) and without (n=44) AAp based on histopathological findings. Twenty-one of the patients without AAp were male and 23 were female. Twenty-one of the patients with AAp were male and 19 were female. The median age of patients without AAp is 43 (95% confidence interval [CI]=35–55) years, and the median age of patients with AAp is 49 (95% CI=41–62) years. No statistically significant difference was found between the groups in terms of categorical variables, except for the imaging variable ($p<0.001$; odds ratio [OR]= 259; 96% CI=41–1636)]. On the other hand, statistically significant differences were found between groups in terms of fecal calprotectin ($p<0.001$), WBC ($p<0.001$), and CRP ($p<0.001$) levels. The results of the statistical analyses between the target variable and qualitative and quantitative

Table 2. The results of the statistical analyzes between the target variable and qualitative data

Variables	Categories	Without AAp, n (%)	With AAp, n (%)	p*
Gender	Male	21 (47.7)	21 (52.5)	0.827
	Female	23 (52.3)	19 (47.5)	
Pregnancy	No	33 (75.0)	34 (85.0)	0.386
	Yes	11 (25.0)	6 (15.0)	
Nausea/vomiting	No	13 (29.6)	12 (30.0)	0.999
	Yes	31 (70.4)	28 (70.0)	
Fever	No	30 (68.2)	21 (52.5)	0.213
	Yes	14 (31.8)	19 (47.5)	
ALT	≤40	31 (70.4)	29 (72.5)	0.999
	>40	13 (29.6)	11 (27.5)	
Imaging	Negative	42 (95.4)	3 (7.5)	<0.001
	Positive	2 (4.6)	37 (92.5)	

*Chi-square test with Yate's correction. ALT: Alanine aminotransferase; AAp: Acute appendicitis.

data are given in Tables 2 and 3, respectively.

According to the results obtained from the analyses made between the qualitative input variables and the target variable, no statistically significant relationship was found between the other variables except the imaging variable and the categories of the target variable, AAp and non-AAp. The results of the statistical analyses between the target variable and quantitative data are given in Table 3.

According to the results obtained from the analyses made between the quantitative input variables and the target variable, there was a statistical difference between the categories of the target variable, AAp, and non-AAp, in terms of calprotectin, WBC, CRP variables. At the same time, there was no difference in age and onset to visit variables. The values of the performance criteria of the RF model used in this study to classify the AAp are given in Table 4.

In the training stage, ACC, BC, sensitivity, specificity, PPV, NPV, and F1 score for the RF model were 95.6%, 95.5%, 93.9%, 97.1%, 96.9%, 94.4%, and 95.4%, respectively. Furthermore, in the testing stage, ACC, BC, sensitivity, specificity, PPV, NPV, and F1 scores from the RF model were 93.8%, 93.8%, 87.5%, 100%, 100%, 88.9%, and 93.3%, respectively. In the Figure 1, performance metrics are plotted for the constructed RF model. The graph of the variables associated with the output variable according to the variable importance obtained from the modeling is given in Figure 2.

DISCUSSION

In this case–control study, predictive variables associated with the diagnosis of AAp were revealed by using RF, which is an ML method. In addition, which of these factors is more important was determined by the variable significance values obtained as a result of the RF modeling.

The ideal approach in the treatment of AAp is to make a correct diagnosis and plan rapid treatment as soon as possible. However, in practice, it is not always possible to achieve this expectation. Because it is relatively easier to diagnose

Table 3. The results of the statistical analyses between the target variable and quantitative data

Variables	Median (95% CI)		p*
	Without AAp	With AAp	
Fecal calprotectin (μg/g)	61 (44–80)	259 (221–298)	<0.001
WBC	8.535 (7.5–9.5)	17.27 (14.2–21.2)	<0.001
CRP	8.92 (7.6–21.2)	32.45 (20.6–38.7)	<0.001
Age	43 (35–55)	49 (41–62)	0.531
Onset to visit	27.5 (20–41)	29 (15–38)	0.792

*Mann–Whitney U-test; FC: Fecal calprotectin; WBC: White blood cell; CRP: C-reactive protein; CI: Confidence interval; AAp: Acute appendicitis.

Table 4. Performance metrics of the random forest model

Metric	Percentage (95% CI)	
	Training set value	Test set value
Accuracy	95.6 (90.7–100)	93.8 (81.9–100)
BC	95.5 (90.6–1)	93.8 (81.9–100)
Sensitivity	93.9 (79.8–99.3)	87.5 (47.35–99.7)
Specificity	97.1 (85.1–99.9)	100 (63.10–100)
PPV	96.9 (83.8–99.9)	100
NPV	94.4 (81.3–99.3)	88.9 (51.8–99.7)
F1 score	95.4 (90.4–100)	93.3 (81.1–100)

PPV: Positive predictive value, NPV: Negative predictive value, CI: Confidence interval, BC: Balanced accuracy.

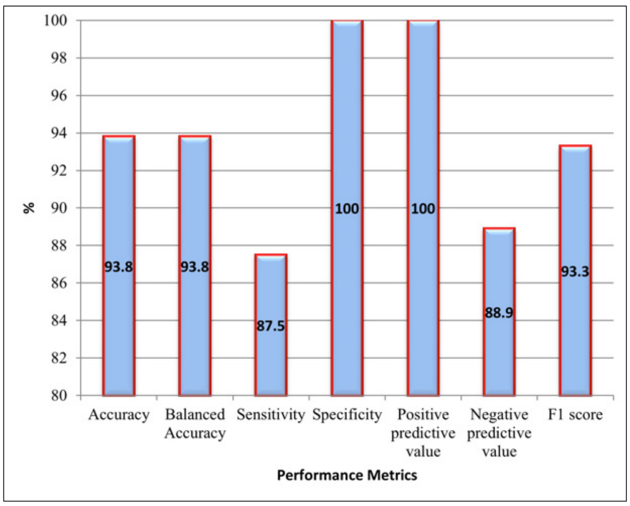

Figure 1. Graph of values for performance metrics for random forest model.

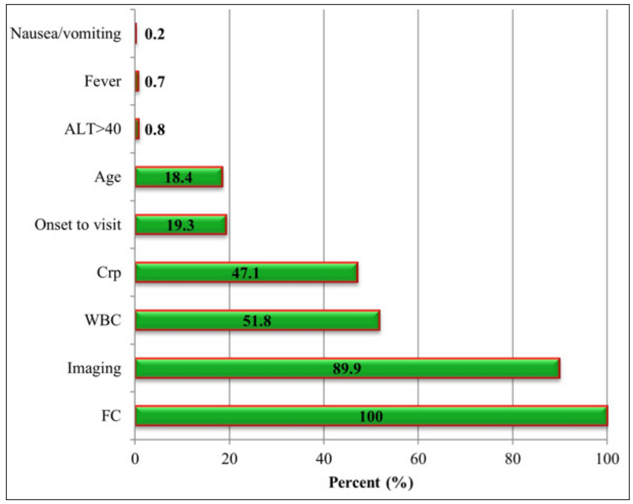

Figure 2. Variable importance graph.

AAp in the emergency services with classical signs and symptoms for patients. On the other hand, some of the patients have atypical symptoms and the diagnostic dilemma in these patients may cause either unnecessary laparotomy (negative appendectomy) or complications due to delay in treatment (perforation, plastron, abscess, and fecal peritonitis).[7]

Negative appendectomy is a condition characterized by the histopathological absence of inflammatory cell infiltration in the appendectomy specimen. This is almost always due to the insufficient and inaccurate use of diagnostic instruments such as biochemical and radiological tools. The main reason for this is that many inflammatory diseases involving the abdominal cavity mimic the signs and symptoms of AAp.[32] Although negative appendectomy rates vary between 19% and 33.9%, they have recently been reduced to <10% in parallel with the more frequent and meticulous use of diagnostic instruments.[33-35] The most important complication associated with AAp is perforation and other related complications. Perforation is most directly related to delayed admission to the hospital, immunosuppressive conditions, prolongation of the diagnostic processes, or misdiagnosis. This complication is associated with serious morbidity and mortality, especially in the elderly.

As it can be understood from our explanations above, it is still difficult to diagnose accurately and quickly in patients with pre-diagnosis of AAp who applied to the emergency services even today. Therefore, it has become imperative to develop easily accessible, economical, and proper instruments that can help in the differential diagnosis of patients with suspected AAp or complicated AAp, and that these instruments be independent of individuals.

ML methods have been commonly utilized in the prediction, diagnosis, and medical decision support systems of diseases in recent years. With ML methods, which are frequently used in the field of health, early diagnosis of diseases and revealing the factors affecting the disease are carried out.[17,36]

The a ACC, BC, sensitivity, specificity, PPV and NPV, and F1 score metrics obtained with the RF model were 93.8%, 93.8%, 87.5%, 100%, 100%, 88.9%, and 93.3%, respectively. In the metrics of the proposed model, specificity and PPV were 100%, which indicated this model performs well in the diagnosing the AA. However, NPV and sensitivity demonstrated weak predictions as compared to specificity and PPV. Regarding the variable importance values as a result of the model, the variables most associated with the diagnosis of AAp were fecal calprotectin, imaging, WBC, CRP, from symptoms onset to hospital visit, age, ALT >40, fever, and nausea/vomiting, respectively. This ML-based study showed that the two most important independent variables associated with the diagnosis of AAp were fecal calprotectin and imaging modalities, respectively. Fecal calprotectin is a calcium- and zinc-binding heterodimer with a molecular weight of 36.5 kDa identified as the predominant cytosolic protein in neutrophil granulocytes.[37] Ambe et al.[38] demonstrated that the histopathological expression of fecal calprotectin in the appendectomy specimen obtained from patients with AAp was significantly higher than specimens obtained from patients without AAp, and the authors suggested that this situation was directly related to the increase in fecal calprotectin levels. In addition, in medical practice, implementation of FC may be the possible difficulty and time delay in achieving stool specimens of the patients from emergency departments. In the current study, the stool specimens were collected from the emergency department. Fecal calprotectin is resistant to proteolytic degradation and can be stable in feces samples for up to 7 days at room temperature, according to previous research.[26]

Zhou et al.,[26] who have the open data set we used in this study, stated that they determined an optimal threshold value for fecal calprotectin by the ROC curve and that this cutoff point is an effective predictor for AAp. Cikot et al.[39] reported that plasma calprotectin level is markedly increased in patients with AAp and is an excellent tool to be used in the differential diagnosis of patients with and without complicated AAp. When these results are evaluated together, it is understood that both plasma and stool calprotectin levels can be used successfully in the diagnosis of AAp patients.

In a study conducted to identify individuals with AAp, an ACC of 95.31% was obtained with gradient boosted trees, one of the ML methods, and there is no fecal calprotectin among the data that can be used in the diagnosis of AAp in the study.[40] Another study aimed to determine the diagnosis of AAp by using different ML methods and obtained high results. However, in the present study, there is no fecal calprotectin among the parameters that may be associated with the diagnosis.[41] In a recent study, many ML models were used for AAp diagnosis, and the highest ACC result was obtained with the RF model, which is 84%. Similarly, when the data set used in the study is examined, the fecal calprotectin variable is not among the data.[42]

When the studies on AAp in the literature were examined, there was no study that detects fecal calprotectin as a biomarker with ML methods. Therefore, this study is the first to detect fecal calprotectin as an important biomarker using the ML method. According to the results of this study, which was carried out using the ML model, it was concluded that another important instrument that can be used in the differential diagnosis of AAp is the imaging method. Although it has been observed over the last two decades that CT can raise the specificity of diagnostic assessment and lower the probability of negative appendectomy, it has drawbacks such as misunderstanding, radiation exposure, and contrast material.[26] In addition, the necessary equipment and qualified personnel for CT are not always available in the emergency departments of hospitals and especially on night shifts. These problems can be solved by using teleconsultation and remote access methods. Especially in cases where the diagnosis of AAp cannot be clarified by scoring systems such as the Alvarado score, which consists of clinical and biochemical blood parameters, the only instrument to be applied is non-operator dependent CT. Conventional statistical analysis methods have shown that instruments such as US and CT are valuable in the diagnosis of AAp, both in the original text of the study from which the data in this study were achieved and in other clinical studies. However, to our knowledge, there are limited studies investigating the importance of radiological examinations in the prediction of AAp using ML methods.

In a study, ML models were used to determine whether CT is beneficial for the diagnosis of appendicitis in the diagnosis of AAp. As a result of the ML models used in the study, it has been shown that imaging can be used as an aid in the diagnosis of AAp and has better sensitivity and specificity than the Alvarado score.[43]

This study has several limitations. First, the dataset used in the study is an open-access dataset and was obtained from a single center. To generalize the results of the study, results of this study should be supported by further multicenter studies. Second, fecal specimens for fecal calprotectin, which is an important predictor for AAp are potentially challenging to collect in emergency departments. For this reason, to use the recommended fecal calprotectin in diagnosis, the importance of systems that can perform analysis with a small amount of sample without loss of time by improving each step of sample collection is increasing, and the development of such systems should be encouraged. Third, in the open data set used in this study, it is an important limiting factor that there is no clear information about whether US and CT are used individually or together in radiological diagnosis. Therefore, Zhou et al.[26] should have specified which radiological instrument was positive rather than the terms of imaging positive.

Conclusion

With the ML model in this study, a prediction model was developed to diagnose AAp and applied to existing data to improve and assist clinicians' diagnostic decisions. The results obtained revealed effective clinical variables that can predict AAp with high ACC. Owing to the variables determined as a result of the modeling, the decision-making process of the clinicians for the diagnosis of AAp will be easier. Thus, proper and correct diagnosis will help decrease the risks of perforation or unnecessary operations and advance overall outcomes (morbidity, mortality, cost, return to work, and quality of life).

Ethics Committee Approval: This study was approved by the Inonu University Faculty of Medicine Clinical Research Ethics Committee (Date: 06-09-2022, Decision No: 2022/3768).

Peer-review: Externally peer-reviewed.

Authorship Contributions: Concept: Z.K., S.A., C.C.; Design: Z.K., S.A., C.C.; Supervision: Z.K., S.A., C.C.; Materials: Z.K., S.A., C.C.; Data: Z.K., S.A., C.C.; Analysis: Z.K., S.A., C.C.; Literature search: Z.K., S.A., C.C.; Writing: Z.K., S.A., C.C.; Critical revision: Z.K., S.A., C.C.

Conflict of Interest: None declared.

Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES

1. Akbulut S, Bahce ZS, Oztas T, Gumus S, Sogutcu N, Sakarya H, et al. Assessment of demographic, clinical and histopathological features of patients who underwent appendectomy due to a presumed diagnosis of acute appendicitis. Ulus Travma Acil Cerrahi Derg 2021;27:315–24.
2. Sarici KB, Akbulut S, Koc C, Tuncer A, Yılmaz S. Liver transplant versus non-liver transplant patients underwent appendectomy with presumed diagnosis of acute appendicitis: Case-control study. Ulus Travma Acil Cerrahi Derg 2020;26:705–12. [CrossRef]
3. Wu J, Jiang H, Li S, Wu X, Wang P, Sawyer R, et al. Optimising the treatment for uncomplicated acute appendicitis (OPTIMA trial): A protocol for a multicentre, randomised, double-blinded placebo-controlled study. BMJ Open 2022;12:e057793. [CrossRef]
4. Wickramasinghe DP, Xavier C, Samarasekera DN. The worldwide epidemiology of acute appendicitis: An analysis of the global health data exchange dataset. World J Surg 2021;45:1999–2008. [CrossRef]
5. Ferris M, Quan S, Kaplan BS, Molodecky N, Ball CG, Chernoff GW, et al. The global incidence of appendicitis: A systematic review of population-based studies. Ann Surg 2017;266:237–41. [CrossRef]
6. Ahmed HO, Muhedin R, Boujan A, Aziz AH, Hardi RA, Abdulla AA, et al. A five-year longitudinal observational study in morbidity and mortality of negative appendectomy in Sulaimani teaching Hospital/Kurdistan Region/Iraq. Sci Rep 2020;10:2028. [CrossRef]
7. Su D, Li Q, Zhang T, Veliz P, Chen Y, He K, et al. Prediction of acute appendicitis among patients with undifferentiated abdominal pain at emergency department. BMC Med Res Methodol 2022;22:18. [CrossRef]
8. Akbulut S, Koc C, Sahin TT, Sahin E, Tuncer A, Demyati K, et al. An investigation into the factors predicting acute appendicitis and perforated appendicitis. Ulus Travma Acil Cerrahi Derg 2021;27:434–42. [CrossRef]
9. Jeon BG, Kim HJ, Heo SC. CT scan findings can predict the safety of delayed appendectomy for acute appendicitis. J Gastrointest Surg 2019;23:1856–66. [CrossRef]
10. Yang Z, Sun F, Ai S, Wang J, Guan W, Liu S. Meta-analysis of studies comparing conservative treatment with antibiotics and appendectomy for acute appendicitis in the adult. BMC Surg 2019;19:110. [CrossRef]
11. Akbulut S, Caliskan A, Ekin A, Yagmur Y. Left-sided acute appendicitis with situs inversus totalis: Review of 63 published cases and report of two cases. J Gastrointest Surg 2010;14:1422–8. [CrossRef]
12. Kim J, Kim K, Kim J, Yoo J, Jeong W, Cho S, et al. The learning curve in diagnosing acute appendicitis with emergency sonography among novice emergency medicine residents. J Clin Ultrasound 2018;46:305–10.
13. Galai T, Beloosesky OZ, Scolnik D, Rimon A, Glatstein M. Misdiagnosis of acute appendicitis in children attending the emergency department: The experience of a large, tertiary care pediatric hospital. Eur J Pediatr Surg 2017;27:138–41. [CrossRef]
14. Chang YJ, Chao HC, Kong MS, Hsia SH, Yan DC. Misdiagnosed acute appendicitis in children in the emergency department. Chang Gung Med J 2010;33:551–7.
15. Graff L, Russell J, Seashore J, Tate J, Elwell A, Prete M, et al. False-negative and false-positive errors in abdominal pain evaluation: Failure to diagnose acute appendicitis and unnecessary surgery. Acad Emerg Med 2000;7:1244–55. [CrossRef]
16. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. PeerJ 2019;7:e7702. [CrossRef]
17. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019;6:94–8. [CrossRef]
18. Mijwil MM, Abttan RA. Artificial intelligence: A survey on evolution and future trends. Asian J Appl Sci 2021;9:87–92. [CrossRef]
19. Polikar R. Ensemble Learning. In: Cha Z, Yunqian M, editors. Ensemble Machine Learning. 1st ed. New York: Springer; 2012. p. 1–34. [CrossRef]
20. Shin B, Chokshi FH, Lee T, Choi JD. Classification of Radiology Reports using Neural Attention Models. In: Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN). United States: IEEE; 2017. p. 4363–70. [CrossRef]
21. Luz CF, Vollmer M, Decruyenaere J, Nijsten MW, Glasner C, Sinha B. Machine learning in infection management using routine electronic health records: Tools, techniques, and reporting of future technologies. Clin Microbiol Infect 2020;26:1291–9. [CrossRef]
22. Ekelik H, Altas D. Classification of potential residential buyers by using random forest method taking advantage of digital advertising data. J Res Econ 2019;3:28–45.
23. Bestil Hİ, Guvensan MA. A novel multi-tier filtering architecture and smart SMS box for classification of Turkish short messages. Int J Adv Eng Pure Sci 2019;31:17–28. [CrossRef]
24. Korkmaz D, Çelik HE, Kapar M. Botnet detection by using classification and regression trees with random forest algorithms: Example of Van Yüzüncü Yil University. Yuzuncu Yil Univ J Inst Nat Appl Sci 2018;23:297–307.
25. Erdem F, Derinpinar MA, Nasirzadehdizaji R, Selen O, Seker DZ, Bayram B. Coastline extraction by using random forest method; a case study of Istanbul. J Geomatics 2018;3:100–7.
26. Zhou W, Qiao H, Yuan W, Yao Z, Liu K, Wang J, et al. Diagnostic utility of fecal calprotectin in patients presenting to the emergency department with suspected acute appendicitis. Am J Emerg Med 2020;38:2303–7.
27. Breiman L. Random forests. Machine Learning 2001;45:5–32. [CrossRef]
28. Cutler A, Cutler DR, Stevens JR. Random forests. In: Cha Z, Yunqian M, editors. Ensemble Machine Learning. 1st edition. New York: Springer; 2012. p. 157–75. [CrossRef]
29. Mather PM, Koch M. Computer Processing of Remotely-sensed Images: An Introduction. 4th edition. United States: John Wiley and Sons; 2011.
30. Pal M. Random forest classifier for remote sensing classification. Int J Remote Sens 2005;26:217–22. [CrossRef]
31. Arslan AK, Yasar S, Colak C, Yologlu S. WSSPAS: An interactive web application for sample size and power analysis with R using shiny. Turk Klin J Biostat 2018;10:224–46. [CrossRef]
32. Rosen MP, Ding A, Blake MA, Baker ME, Cash BD, Fidler JL, et al. ACR Appropriateness Criteria® right lower quadrant pain--suspected appendicitis. J Am Coll Radiol 2011;8:749–55. [CrossRef]
33. Akbar F, Yousuf M, Morgan RJ, Maw A. Changing management of suspected appendicitis in the laparoscopic era. Ann R Coll Surg Engl 2010;92:65–8. [CrossRef]
34. Lim J, Pang Q, Alexander R. One year negative appendicectomy rates at a district general hospital: A retrospective cohort study. Int J Surg 2016;31:1–4. [CrossRef]
35. Noureldin K, Hatim Ali AA, Issa M, Shah H, Ayantunde B, Ayantunde A. Negative appendicectomy rate: Incidence and predictors. Cureus 2022;14:e21489. [CrossRef]
36. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018;2:719–31. [CrossRef]
37. Ricciuto A, Griffiths AM. Clinical value of fecal calprotectin. Crit Rev Clin Lab Sci 2019;56:307–20. [CrossRef]

38. Ambe PC, Godde D, Bonicke L, Papadakis M, Storkel S, Zirngibl H. Calprotectin could be a potential biomarker for acute appendicitis. J Transl Med 2016;14:107. [CrossRef]
39. Cikot M, Peker KD, Bozkurt MA, Kocatas A, Kones O, Binboga S, et al. Plasma calprotectin level: Usage in distinction of uncomplicated from complicated acute appendicitis. World J Emerg Surg 2016;11:7. [CrossRef]
40. Akmese OF, Dogan G, Kor H, Erbay H, Demir E. The use of machine learning approaches for the diagnosis of acute appendicitis. Emerg Med Int 2020;2020:7306435. [CrossRef]
41. Aydin E, Turkmen İU, Namli G, Ozturk C, Esen AB, Eray YN, et al. A novel and simple machine learning algorithm for preoperative diagnosis of acute appendicitis in children. Pediatr Surg Int 2020;36:735–42.
42. Mijwil MM, Aggarwal K. A diagnostic testing for people with appendicitis using machine learning techniques. Multimed Tools Appl 2022;81:7011–23. [CrossRef]
43. Gunasingha RM, Grey SF, Munoz B, Schobel S, Lee J, Erwin C, et al. To scan or not to scan: Development of a clinical decision support tool to determine if imaging would aid in the diagnosis of appendicitis. World J Surg 2021;45:3056–64. [CrossRef]

ORİJİNAL ÇALIŞMA - *ÖZ*

Önerilen bir makine öğrenimi modeline dayalı akut apandisit öngörüsünde fekal kalprotektinin değeri

Dr. Zeynep Küçükakçali,[1] Dr. Sami Akbulut,[2] Dr. Cemil Çolak[1]

[1]İnönü Üniversitesi Tıp Fakültesi, Biyoistatistik ve Medikal İnformatik Anabilim Dalı, Malatya, Türkiye

[2]İnönü Üniversitesi Tıp Fakültesi, Genel Cerrahi Anabilim Dalı, Malatya, Türkiye

AMAÇ: Bu çalışmanın amacı, makine öğrenmesi (ML) algoritmalarından biri olan Random Forest'i (RF) akut apandisit (AAp) ön tanısı olan hastalardan oluşan bir veri setine uygulamak ve AAp tanısı ile ilişkili en önemli faktörleri değişken önemliliğine göre ortaya koymaktır.

GEREÇ VE YÖNTEM: Bu vaka-kontrol çalışmasında AAp için biyobelirteçleri tahmin etmek üzere AAp'si olan (n=40) ve olmayan (n=44) iki hasta grubunu karşılaştıran açık erişimli bir veri seti kullanıldı. Veri setinin modellenmesinde RF kullanıldı. Veriler eğitim ve test veri seti olmak üzere ikiye ayrıldı (80: 20). Model performansı için doğruluk, dengeli doğruluk, duyarlılık, özgüllük, pozitif tahmin değeri (PPV) ve negatif tahmin değeri (NPV) performans metrikleri değerlendirildi.

BULGULAR: RF modeline ait doğruluk, dengeli doğruluk, duyarlılık, özgüllük, PPV, NPV ve F1 skorları sırasıyla %93.8, %93.8, %87.5, %100, %100, %88.9 ve %93.3 olarak hesaplandı. Modele ilişkin değişken önem değerlerinin ardından, AAp tanısı ve öngürüsü ile en çok ilişkili olan değişkenler sırasıyla fekal kalprotektin (%100), radyolojik görüntüleme (%89,9), beyaz kan testi (%51,8), C-reaktif protein (%47,1), semptomların hastane ziyaretinde başlaması (%19.3), hasta yaşı (%18.4), ALT düzeyleri >40 (<%1), ateş (< %1) ve mide bulantısı/kusma (<%1) olarak belirlendi.

TARTIŞMA: Bu çalışmada ML yöntemi ile AAp için bir tahmin modeli geliştirildi. Bu model sayesinde AAp'i yüksek doğruluk ile öngören biyobelirteçler belirlendi. Böylece klinisyenlerin AAp tanısına karar verme süreci kolaylaşacak, yüksek doğruluk ile zamanında tanı konulabilmesi sayesinde perforasyon ve gereksiz ameliyat riskleri en aza indirilecektir.

Anahtar sözcükler: Akut apandisit; makine öğrenimi; random forest; modelleme; değişken önemliliği; fekal kalprotektin.

Ulus Travma Acil Cerrahi Derg 2023;29(6):655-662 doi: 10.14744/tjtes.2023.10001