Inflammatory indices, machine learning and artificial intelligence in tubal ectopic pregnancy management

Tubal ektopik gebelik yönetiminde enflamatuvar indeksler, makine öğrenmesi ve yapay zeka

 Uğurcan Zorlu[1], Senem Arda Düz[2], Gül Kurtaran[3], Mohammad İbrahim Halilzade[3], Burak Elmas[4]

[1]Ankara Bilkent City Hospital, Clinic of Perinatology, Ankara, Türkiye
[2]İnönü University Faculty of Medicine, Department of Obstetrics and Gynecology, Malatya, Türkiye
[3]Ankara Bilkent City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Türkiye
[4]Acıbadem Maslak Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Türkiye

Abstract

Objective: To assess the predictive value of hematologic and biochemical inflammatory indices for methotrexate (MTX) treatment outcomes in tubal ectopic pregnancy (TEP) and to develop machine learning (ML) models for individualized risk stratification.

Materials and Methods: This retrospective cohort included 293 hemodynamically stable TEP patients who were treated with a single dose of MTX between January 2019 and December 2023. Demographic, clinical, ultrasonographic, and laboratory data were analyzed. Inflammatory indices—including neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio, systemic immune-inflammation index, systemic inflammation response index (SIRI), aggregate index of systemic inflammation (AISI), and fibrinogen-to-albumin ratio (FAR)—were calculated. Outcomes were categorized as single-dose MTX success, requirement for additional MTX, or surgery. Predictive accuracy of five supervised ML algorithms was evaluated using receiver operating characteristic analysis.

Results: Single-dose MTX was successful in 65.5% of patients; 18.4% required an additional dose, and 16.0% underwent surgery. AISI had the highest predictive accuracy for surgery [area under the curve (AUC)=0.929], followed by SIRI (AUC=0.899) and FAR (AUC=0.847). NLR best predicted the need for additional MTX (AUC=0.675). Na¨ve Bayes achieved the highest performance for surgical prediction (accuracy=98.3%, AUC=0.998), while random forest and gradient boosting were most effective in predicting the need for additional MTX (accuracy=83.1%, AUC=0.884-0.896). Feature importance analyses consistently ranked AISI, SIRI, and FAR as top predictors.

Conclusion: AISI, SIRI, and FAR are strong predictors of MTX failure and surgical intervention in TEP. Combining these biomarkers with ML models markedly improves predictive performance and supports a personalized approach to TEP management. Multicenter prospective validation is needed before clinical application.

Keywords: Ectopic pregnancy, inflammation, inflammatory markers, machine learning, methotrexate

Öz

Amaç: Tubal ektopik gebelikte (TEG) metotreksat (MTX) tedavi sonuçlarını öngörmede hematolojik ve biyokimyasal enflamatuvar indekslerin tahmin gücünü değerlendirmek ve bireyselleştirilmiş risk sınıflaması için makine öğrenmesi (MÖ) modelleri geliştirmektir.

Gereç ve Yöntemler: Bu retrospektif kohort çalışmaya Ocak 2019-Aralık 2023 tarihleri arasında tek doz MTX ile tedavi edilen 293 hemodinamik olarak stabil TEG hastası dahil edildi. Demografik, klinik, ultrasonografik ve laboratuvar verileri analiz edildi. Nötrofil-lenfosit oranı (NLR), trombosit-lenfosit oranı, sistemik immün-enflamasyon indeksi, sistemik enflamasyon yanıt indeksi (SIRI), toplu sistemik enflamasyon indeksi (AISI)

PRECIS: Using inflammatory indices and machine learning models, we evaluated predictors of methotrexate failure and surgical intervention in patients with tubal ectopic pregnancy.

ve fibrinojen/albümin oranı (FAR) gibi enflamatuvar indeksler hesaplandı. Tedavi sonuçları; tek doz MTX başarısı, ek doz MTX gereksinimi ve cerrahi müdahale olarak sınıflandırıldı. Tahmin doğruluğu, ROC analizi ve beş farklı denetimli MÖ algoritması ile değerlendirildi.

Bulgular: Tek doz MTX tedavisi %65,5 oranında başarılı olurken; %18,4 hastada ek doz MTX gereksinimi oluştu, %16,0 hastaya ise cerrahi müdahale uygulandı. Cerrahi tahmininde en yüksek doğruluk AISI'ye ait olup eğrinin altındaki alan (AUC) değeri 0,929 olarak bulundu; bunu SIRI (AUC=0,899) ve FAR (AUC=0,847) izledi. Ek MTX ihtiyacını en iyi öngören indeks NLR oldu (AUC=0,675). Na¨ve Bayes algoritması, cerrahi öngörüsünde en yüksek başarıyı sağladı (doğruluk=%98,3; AUC=0,998); rastgele orman (random forest) ve gradyan artırma (gradient boosting) algoritmaları ise ek MTX ihtiyacını öngörmede en etkili modellerdi (doğruluk=%83,1; AUC=0,884-0,896). Özellik önem analizleri, AISI, SIRI ve FAR indekslerini tutarlı şekilde en güçlü öngörücüler olarak belirledi.

Sonuç: AISI, SIRI ve FAR, TEG'de MTX tedavi başarısızlığını ve cerrahi müdahale gereksinimini öngörmede güçlü biyobelirteçlerdir. Bu biyobelirteçlerin MÖ modelleriyle entegrasyonu, tahmin performansını önemli ölçüde artırarak TEG yönetiminde kişiselleştirilmiş bir yaklaşımı destekler. Klinik uygulamaya geçmeden önce çok merkezli prospektif çalışmalarla doğrulama gereklidir.

Anahtar Kelimeler: Ektopik gebelik, enflamasyon, enflamatuvar belirteçler, makine öğrenmesi, metotreksat

Introduction

Tubal ectopic pregnancy (TEP) remains a significant contributor to maternal morbidity and mortality during the first trimester and constitutes more than 90% of all ectopic pregnancies[1]. The use of systemic methotrexate (MTX) has substantially changed the management of hemodynamically stable patients, as single-dose regimens provide a less invasive alternative to surgical intervention[2]. However, the clinical response to MTX is heterogeneous, and a subset of patients experiences treatment failure, necessitating additional MTX administration or surgical management[3,4]. For this reason, the early identification of patients at increased risk for unsuccessful medical treatment is essential to enable individualized therapeutic strategies and improve overall prognosis. In recent years, hematologic and biochemical inflammatory indices obtained from routine laboratory testing, such as the systemic immune-inflammation index (SII), platelet-to-lymphocyte ratio (PLR), neutrophil-to-lymphocyte ratio (NLR), and fibrinogen-to-albumin ratio (FAR), have been increasingly investigated as potential predictors of disease course and treatment response in obstetric and gynecologic disorders[4,5]. These markers reflect systemic immune activation, coagulation status, and metabolic changes, all of which may influence MTX efficacy in TEP[6,7]. Although several studies have explored their predictive role, findings remain inconsistent, particularly regarding their ability to forecast the need for surgical intervention versus additional MTX administration[8,9].

Alongside biomarker research, advances in artificial intelligence and machine learning (ML) offer novel opportunities to improve clinical decision-making in TEP. ML algorithms can process complex, multidimensional datasets to identify nonlinear relationships between clinical, ultrasonographic, and laboratory variables, potentially outperforming traditional statistical approaches[10,11]. However, integration of ML-based prediction models into routine ectopic pregnancy management remains limited, with few studies systematically comparing their performance against established clinical predictors.

Given these gaps, the present retrospective cohort study aimed to evaluate the predictive accuracy of inflammatory indices for treatment success, the requirement for additional MTX, and surgical intervention in TEP patients, and to develop and validate ML-based prediction models for individualized risk stratification. By combining traditional statistical methods with advanced computational modeling, this study seeks to establish a more precise, data-driven framework for optimizing TEP management.

Materials and Methods

This retrospective cohort study was conducted in the Department of Obstetrics and Gynecology of a tertiary referral hospital between January 2019 and December 2023. All clinical, laboratory, and imaging data were retrieved from the institutional electronic medical record system and radiology archives. The study protocol was reviewed and approved by the Ankara Bilkent City Hospital Institutional Ethics Committee (approval no: TABED 2-25-1311, date: 11.06.2025), and all procedures were conducted in accordance with the Declaration of Helsinki.

All women diagnosed with TEP during the study period were evaluated for eligibility. A definitive diagnosis was established based on transvaginal ultrasonography findings and/or serial serum beta human chorionic gonadotropin (β-hCG) measurements. Only hemodynamically stable patients who were initially managed with a single-dose intramuscular MTX regimen and had complete clinical, laboratory, and imaging records were included. Patients presenting with hemodynamic instability, suspected or confirmed tubal rupture, immediate indications for surgical intervention, contraindications to MTX therapy (e.g., hepatic, renal, or hematologic disorders), non-tubal ectopic pregnancies (including cervical, interstitial, or ovarian locations), or those treated primarily with multi-dose MTX protocols were excluded from the analysis.

All eligible patients received a single intramuscular dose of MTX at 50 mg/m² on day 0. Serum β-hCG levels were measured on days 0, 4, and 7 following treatment. A decline of at least 15% in β-hCG levels between day 4 and day 7 was considered indicative of an adequate therapeutic response. Patients who failed to achieve this decline received an additional dose of MTX and were classified as requiring further medical treatment. Those who developed worsening clinical

Zorlu et al. Inflammation and ML in ectopic pregnancy

Turk J Obstet Gynecol 2026;23(1):47-55

symptoms such as increasing abdominal pain, hemodynamic deterioration, signs of tubal rupture, or persistent elevation of β-hCG levels despite medical therapy were referred for surgical management.

Demographic characteristics including age, gravidity, and parity, as well as clinical features such as abdominal pain and vaginal bleeding at presentation, were recorded for each patient. Laboratory parameters included serum β-hCG levels at baseline, day 4, and day 7; complete blood count values; and concentrations of C-reactive protein, fibrinogen, and albumin. Ultrasonographic evaluation provided data on adnexal mass size and Doppler-derived ipsilateral ovarian artery systolic/diastolic ratio and pulsatility index. Based on complete blood count parameters, several hematologic and inflammatory indices were calculated, including NLR, monocyte-to-lymphocyte ratio (MLR), PLR, eosinophil-to-lymphocyte ratio (ELR), white blood cell-to-neutrophil ratio, SII, systemic inflammation response index (SIRI), aggregate index of systemic inflammation (AISI), and FAR.

The primary outcomes of the study were defined as the successful resolution of ectopic pregnancy with a single dose of methotrexate, the requirement for an additional MTX dose, and the need for surgical intervention following medical treatment. Secondary outcomes included the identification of clinical and laboratory predictors associated with treatment failure, the assessment of the diagnostic performance of inflammatory indices in predicting treatment outcomes, and the evaluation of ML models for predicting surgical intervention or the need for an additional dose of methotrexate.

Statistical Analysis

All analyses were conducted using IBM SPSS Statistics version 25.0 (IBM Corp., Armonk, NY, USA) and Python 3.10 with standard scientific libraries. Data distribution was assessed with the Shapiro-Wilk test. Parametric data were reported as mean ± standard deviation and compared using the Student's t-test or one-way ANOVA, whereas nonparametric variables were expressed as median (interquartile range) and analyzed with the Mann-Whitney U or Kruskal-Wallis tests. Categorical variables were compared using the chi-square test or Fisher's exact test, as appropriate. A two-sided p-value <0.05 was considered statistically significant.

Receiver operating characteristic (ROC) analyses were performed to assess the ability of inflammatory indices to predict surgical intervention and the requirement for an additional dose of methotrexate. Area under the curve (AUC) values and 95% confidence intervals (CIs) were calculated using the DeLong method. Optimal cut-off points were identified with the Youden Index, and sensitivity, specificity, positive predictive value, and negative predictive value were subsequently derived.

Machine Learning Analysis

A supervised ML framework was applied to construct predictive models addressing two binary classification objectives: identifying patients requiring surgical intervention and predicting the need for an additional dose of methotrexate. These models were designed to support individualized risk stratification in the management of tubal ectopic pregnancy. A supervised ML framework was applied using standardized variables, an 80:20 stratified train–test split, and five-fold cross-validation, with performance evaluated using accuracy, AUC, and F1 score.

Prior to model training, comprehensive data preprocessing was undertaken. All continuous clinical, hematologic, and inflammatory variables were standardized using Z-score transformation to eliminate scale-related bias and ensure equal contribution of features during model learning. The proportion of missing data was low, accounting for less than two percent of all entries, and was addressed through mean imputation for continuous variables and mode imputation for categorical variables. The class distribution was examined for imbalance; because surgical intervention outcomes demonstrated moderate skewness, class-weighting strategies were incorporated into algorithms sensitive to class imbalance, such as logistic regression and support vector machines, to enhance model stability and predictive accuracy. For feature selection, all available hematologic and inflammatory indices were initially included as candidate predictors. The relative importance of each variable was subsequently evaluated using two complementary approaches. First, the mean decrease in impurity derived from random forest models was calculated to assess each feature's contribution to reducing classification error. Second, permutation-based importance analysis was performed to quantify the change in model performance following random shuffling of individual features, thereby directly measuring their impact on predictive accuracy. The results of these analyses are presented in Figures 1 and 2.

For predictive modeling, commonly used supervised classification algorithms were applied: logistic regression with L2 regularization, random forest, Gaussian naïve Bayes, support vector machine with a radial basis function kernel, and gradient boosting. Model hyperparameters were optimized using grid search strategies, with the number of trees in the random forest set to 500 and the optimal tree depth determined algorithmically. For the support vector machine model, the penalty parameter was tuned to achieve optimal classification performance, while gradient boosting models were trained using a learning rate of 0.1, with depth parameters refined through cross-validation.

To ensure robust evaluation, the dataset was randomly divided into training and testing subsets at an 80:20 ratio, stratified according to outcome categories, and a fixed random seed was applied to guarantee reproducibility. Hyperparameter tuning was conducted using five-fold cross-validation within
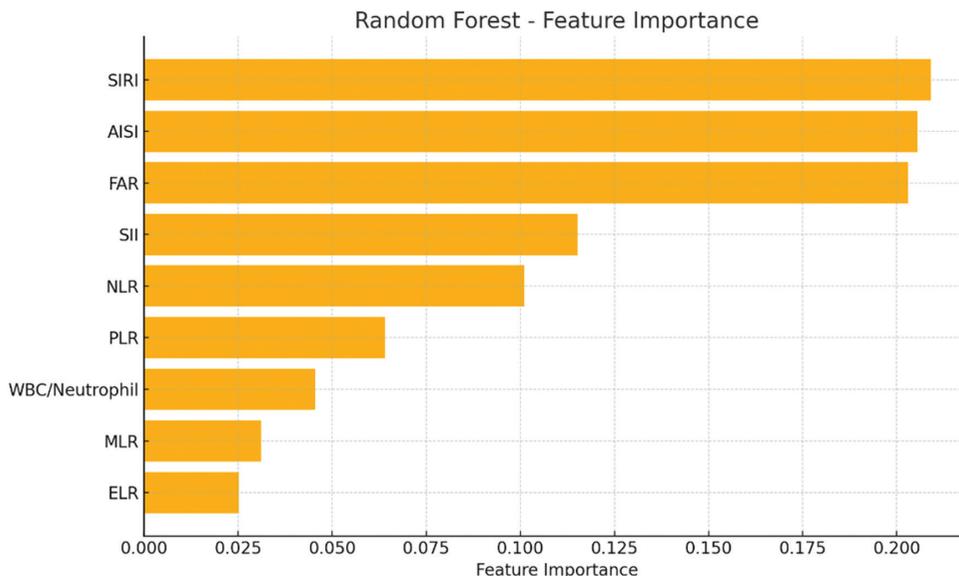
Random Forest - Feature Importance

Figure 1. Top predictors of surgical requirement (random forest feature importance)

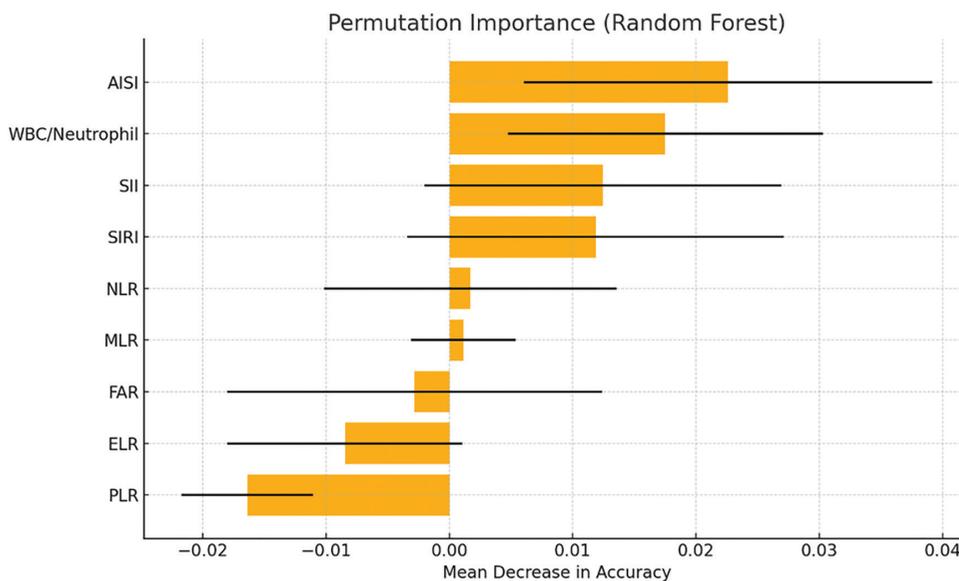## Permutation Importance (Random Forest)

Figure 2. Permutation-based feature importance for predicting surgical intervention in ectopic pregnancy using random forest classifier

the training dataset. Final model performance was assessed on the independent test dataset. For feature-importance analyses and graphical representations, models were subsequently retrained on the entire dataset to maximize statistical power and enhance generalizability.

Model performance was quantified using multiple complementary metrics. Overall classification accuracy was calculated to determine the proportion of correctly classified cases. Discriminative capacity, independent of decision thresholds, was assessed using the area under the ROC curve. The F1 score was employed to capture the balance between precision and recall, and confusion matrices were generated

to examine the distribution of false-positive and false-negative predictions.

To identify the most influential variables associated with surgical intervention, feature importance analyses were performed using the final random forest model trained on the complete dataset. Variable contributions were first assessed using mean decrease in impurity, which reflects the relative importance of each feature in reducing classification error across decision trees. In addition, permutation-based importance analysis was conducted by randomly shuffling individual features and quantifying the resulting decline in model accuracy, thereby directly measuring the dependence

Zorlu et al. Inflammation and ML in ectopic pregnancy

Turk J Obstet Gynecol 2026;23(1):47-55

of predictive performance on each variable. Both methods were applied to ensure consistency and robustness of feature ranking. Across both approaches, aggregate index of systemic inflammation, systemic inflammation response index, and FAR consistently emerged as the most influential predictors, followed by neutrophil-to-lymphocyte and PLR ratios. All machine-learning analyses and visualizations were performed using the scikit-learn library and Matplotlib.

Results

A total of 293 women diagnosed with TEP between January 2019 and December 2023 were included in this retrospective study. Among these patients, 192 (65.5%) achieved complete resolution with a single intramuscular dose of methotrexate, 54 (18.4%) required an additional dose of medical therapy, and 47 (16.0%) ultimately underwent surgical management.

Clinical Characteristics

A comparative analysis of clinical features by treatment outcome is presented in Table 1. The mean age did not differ significantly among the single-dose MTX group, the additional-dose group, and the surgical intervention group (p=0.195). Baseline serum β-hCG concentrations, adnexal mass dimensions, and gestational age at diagnosis were also comparable across the three groups, with no statistically significant differences observed (p=0.552, p=0.376, and p=0.203, respectively).

Vaginal bleeding at presentation was observed more frequently in patients who required an additional dose of MTX than in those who responded to a single dose or who proceeded to surgery; this difference was statistically significant (p=0.038). In contrast, the occurrence of abdominal pain was similar across all groups and was not significantly associated with treatment outcome (p=0.616).

Doppler ultrasonographic findings revealed marked differences between groups. The ipsilateral ovarian artery systolic/diastolic ratio was significantly higher in patients who underwent surgery than in those who were successfully treated with a single dose of MTX or who required an additional dose of MTX (p<0.001). Likewise, the pulsatility index was substantially elevated in the surgical group compared with both medical treatment groups, indicating increased vascular resistance in patients progressing to surgical management (p<0.001).

Hematologic and Inflammatory Parameters

Detailed comparisons of hematologic and inflammatory indices are provided in Table 2. The NLR demonstrated a stepwise increase across outcome groups, being lowest in patients successfully treated with a single dose of methotrexate, higher in those requiring an additional dose, and highest among patients undergoing surgical intervention (p<0.001).

A similar pattern was observed for the MLR, PLR, and eosinophil-to-lymphocyte ratio, all of which were significantly elevated in patients who required further medical treatment or surgery, compared with those who responded to initial therapy (p<0.01 for MLR and PLR; p=0.048 for ELR). Systemic inflammatory markers, including the systemic immune-inflammation index, systemic inflammation response index, aggregate index of systemic inflammation, and fibrinogen-to-albumin ratio, were also significantly higher in the additional-dose and surgical groups than in the single-dose group (all p<0.05). Consistent with these findings, C-reactive protein levels increased progressively across groups, with the highest values observed in patients who underwent surgery (p=0.002). In contrast, the white blood cell-to-neutrophil ratio was significantly lower in the surgical intervention group than in patients successfully treated with a single dose of methotrexate, suggesting a shift toward neutrophil predominance in more severe disease (p=0.032).

ROC Analysis for Predicting Surgical Intervention

The ROC analysis results for inflammatory markers predicting surgical intervention are presented in Table 3. The AISI demonstrated the highest discriminatory ability, with an AUC of 0.929 (95% CI: 0.892–0.963) at a cut-off value of

Table 1. Clinical parameters by treatment outcome in patients with ectopic pregnancy (mean ± SD or %)

Parameter	Single-dose MTX	Additional-dose MTX	Surgical intervention	p-value
Age (years)	28.00±6.02	27.83±5.46	26.27±6.03	0.195
Serum β-hCG (beta-human chorionic gonadotropin, IU/L)	3263.72±1210.15	3418.07±1163.36	3433.58±1237.33	0.552
Adnexal mass size (cm)	2.79±0.70	2.93±0.61	2.87±0.63	0.376
Gestational age (weeks)	5.60±0.31	5.55±0.20	5.65±0.28	0.203
Vaginal bleeding (presence, %)	56.8%	75.9%	59.6%	0.038
Abdominal pain (presence, %)	68.8%	74.1%	74.5%	0.616
Ipsilateral ovarian artery S/D	3.14±0.52	3.53±0.39	4.03±0.53	<0.001
Ipsilateral ovarian artery PI	1.11±0.20	1.32±0.20	1.61±0.24	<0.001

MTX: Methotrexate, β-hCG: Beta-human chorionic gonadotropin, SD: Standard deviation, S/D: Systolic/diastolic ratio, PI: Pulsatility index

221.678, yielding a sensitivity of 95.7% and a specificity of 77.2%. This was followed by the systemic inflammation response index (SIRI) (AUC=0.899, 95% CI: 0.846-0.944; cut-off 738.615; sensitivity 72.3%; specificity 95.5%) and the fibrinogen-to-albumin ratio (FAR) (AUC=0.847, 95% CI: 0.761-0.920; cut-off 0.129; sensitivity 76.6%; specificity 80.5%).

Other markers such as NLR, MLR, PLR, and SII also demonstrated good discriminatory performance (AUC range: 0.793–0.835), whereas the white blood cell to neutrophil ratio showed poor predictive value (AUC=0.241).

Model Performance for Surgical Intervention

The performances of ML models using inflammatory markers as predictors are summarized in Table 4. The Naive Bayes (NB) model achieved the highest overall performance, with an accuracy of 98.3%, an ROC AUC of 0.998, and an F1 score of 0.982.

Logistic Regression (accuracy 96.6%, ROC AUC 0.996, F1=0.964) and Support Vector Machine (accuracy 94.9%, ROC AUC 0.996, F1=0.948) also demonstrated excellent predictive performance. Random Forest and Gradient Boosting yielded slightly lower accuracies (93.2% each) but maintained high ROC AUC values (0.991 and 0.989, respectively).

Feature importance rankings from the Random Forest analysis are illustrated in Figure 1, highlighting AISI, SIRI, and FAR as the most influential predictors of surgical requirement. Figure 2 depicts permutation-based feature importance, confirming the predominance of these indices.

ROC Analysis for Predicting Additional-dose MTX Requirement

The ROC analysis for predicting additional-dose MTX is shown in Table 5.

Table 2. Comparison of hematologic indices by treatment outcome in patients with tubal ectopic pregnancy (mean ± SD)

Parameter	Single-dose MTX	Additional-dose MTX	Surgical intervention	p-value
NLR	2.91±1.00	4.33±1.19	4.98±1.87	<0.001
MLR	0.26±0.08	0.32±0.09	0.38±0.09	<0.01
PLR	115.55±28.22	139.51±32.60	162.10±41.02	<0.01
ELR	0.042±0.020	0.063±0.030	0.067±0.026	0.048
WBC/neutrophil ratio	1.60±0.31	1.50±0.28	1.33±0.19	0.032
SII (×10^9/L)	0.83±0.32	1.05±0.41	1.34±0.50	<0.001
SIRI	526.33±108.91	661.72±133.04	789.45±146.37	<0.001
AISI	152.80±49.90	212.75±64.35	279.93±73.12	<0.001
FAR	0.10±0.02	0.12±0.02	0.14±0.03	0.017
CRP (mg/L)	4.2±1.1	6.5±1.3	9.8±2.1	0.002

MTX: Methotrexate, NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, ELR: Eosinophil-to-lymphocyte ratio, WBC: White blood cell, SII: Systemic immune-inflammation index, SIRI: Systemic inflammation response index, AISI: Aggregate index of systemic inflammation, FAR: Fibrinogen-to-albumin ratio, CRP: C-reactive protein, SD: Standard deviation

Table 3. ROC analysis of inflammatory markers for predicting surgical intervention in patients with tubal ectopic pregnancy using the Youden Index method

Parameter	AUC	95% CI	Cut-off (Youden)	Sensitivity	Specificity
NLR	0.800	0.725-0.877	4.360	0.638	0.841
MLR	0.764	0.699-0.830	0.338	0.787	0.659
PLR	0.793	0.705-0.869	145.869	0.681	0.797
ELR	0.700	0.618-0.774	0.0646	0.532	0.785
WBC/neutrophil ratio	0.241	0.186-0.295	1.191	0.979	0.057
SII	0.835	0.759-0.900	1.14	0.830	0.728
SIRI	0.899	0.846-0.944	738.615	0.723	0.955
AISI	0.929	0.892-0.963	221.678	0.957	0.772
FAR	0.847	0.761-0.920	0.129	0.766	0.805

NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, ELR: Eosinophil-to-lymphocyte ratio, WBC: White blood cell, SII: Systemic immune-inflammation index, SIRI: Systemic inflammation response index, AISI: Aggregate index of systemic inflammation, FAR: Fibrinogen-to-albumin ratio, AUC: Area under the curve, CI: Confidence interval, ROC: Receiver operating characteristic

Zorlu et al. Inflammation and ML in ectopic pregnancy

Turk J Obstet Gynecol 2026;23(1):47-55

Among the evaluated markers, NLR exhibited the highest predictive capacity (AUC=0.675, 95% CI: 0.603-0.747) with an optimal cut-off value of 3.842 (sensitivity 90.7%, specificity 43.5%). PLR (AUC=0.642) and ELR (AUC=0.626) provided moderate discrimination, whereas WBC/Neutrophil ratio (AUC=0.435) demonstrated poor predictive value.

Machine Learning Model Performance for Additional-dose MTX Prediction

As presented in Table 6, the best-performing models for predicting additional-dose MTX requirement were random forest and gradient boosting. Both models achieved an accuracy of 83.1%; their ROC AUC values were 0.884 and 0.896 for random forest and gradient boosting, respectively. A Support Vector Machine achieved an accuracy of 81.4% with an ROC AUC of 0.809. In contrast, Logistic Regression and Naive Bayes models achieved moderate accuracy (72.9%) and similar ROC AUC values (0.805 and 0.807, respectively).

Table 4. Machine learning model performance for predicting surgical intervention in patients with tubal ectopic pregnancy using NLR, MLR, PLR, ELR, WBC/Neutrophil ratio, SII, SIRI, AISI, and FAR

Model	Accuracy	ROC-AUC	F1 score
Logistic regression	0.966	0.996	0.964
Random forest	0.932	0.991	0.931
Naive bayes	0.983	0.998	0.982
Support vector machine	0.949	0.996	0.948
Gradient boosting	0.932	0.989	0.930

NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, ELR: Eosinophil-to-lymphocyte ratio, WBC: White blood cell, SII: Systemic immune-inflammation index, SIRI: Systemic inflammation response index, AISI: Aggregate index of systemic inflammation, FAR: Fibrinogen-to-albumin ratio, ROC-AUC: Receiver operating characteristic-area under the curve, F1 Score: Harmonic mean of precision and recall

Table 5. ROC analysis of hematologic and inflammatory markers for predicting additional-dose MTX requirement in patients with tubal ectopic pregnancy using the Youden Index method

Parameter	AUC	95% CI	Cut-off (Youden)	Sensitivity	Specificity
NLR	0.675	0.603-0.747	3.842	0.907	0.435
MLR	0.557	0.477-0.630	0.296	0.722	0.427
PLR	0.642	0.563-0.717	129.492	0.815	0.448
ELR	0.626	0.530-0.719	0.058	0.611	0.686
WBC/neutrophil ratio	0.435	0.360-0.515	1.219	0.019	0.996
SII	0.590	0.505-0.671	0.994	0.574	0.624
SIRI	0.611	0.523-0.692	621.223	0.667	0.573
AISI	0.602	0.520-0.682	190.456	0.630	0.585
FAR	0.578	0.494-0.658	0.112	0.537	0.616

MTX: Methotrexate, NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, ELR: Eosinophil-to-lymphocyte ratio, WBC: White blood cell, SII: Systemic immune-inflammation index, SIRI: Systemic inflammation response index, AISI: Aggregate index of systemic inflammation, FAR: Fibrinogen-to-albumin ratio, AUC: Area under the curve, CI: Confidence interval

Table 6. Performance metrics of machine learning models for predicting the need for additional-dose methotrexate (MTX) in patients with tubal ectopic pregnancy

Model	Accuracy	ROC-AUC	F1 score
Logistic regression	0.729	0.805	0.511
Random forest	0.831	0.884	0.545
Naive bayes	0.729	0.807	0.513
Support vector machine	0.814	0.809	0.521
Gradient boosting	0.831	0.896	0.543

MTX: Methotrexate, ROC-AUC: Receiver operating characteristic-area under the curve, F1 Score: Harmonic mean of precision and recall

Turk J Obstet Gynecol 2026;23(1):47-55

Zorlu et al. Inflammation and ML in ectopic pregnancy

Discussion

This study provides robust evidence that specific inflammatory indices, particularly AISI, SIRI, and FAR, strongly predict MTX treatment failure and the need for surgical intervention in TEP. Furthermore, our results demonstrate that ML algorithms, especially Naïve Bayes and logistic regression, can achieve excellent predictive accuracy, outperforming traditional cut-off-based approaches.

Our findings are in line with recent studies investigating the prognostic utility of hematologic markers in ectopic pregnancy. Bilir et al.[8] reported that hemogram-based indices such as NLR, PLR, and SII were significantly higher in patients requiring surgical management after MTX therapy, with AISI emerging as the most discriminatory parameter. Dinc and Issın[5] showed that elevated SII values at presentation correlated strongly with the risk of tubal rupture, underscoring the role of systemic inflammation in the progression of ectopic pregnancy. Seyfettinoglu and Adiguzel[9] further highlighted that combining multiple indices improved predictive performance compared to single parameters.

The application of ML in ectopic pregnancy prognosis remains nascent but shows promise. Chen et al.[11] used gradient boosting and random forest models to predict MTX success in ectopic pregnancy and found ROC-AUC values up to 0.94-comparable to the performance observed in our Naïve Bayes model (AUC=0.998). Our inclusion of Doppler ultrasonography parameters alongside inflammatory indices may partly explain the superior performance, as ultrasonographic vascular indices reflect tubal perfusion and inflammatory status. ROC-derived cut-off values should be interpreted as supportive risk indicators rather than absolute clinical thresholds and must be integrated with clinical assessment and β-hCG dynamics.

Clinically, integrating these biomarkers with ML-based prediction tools could facilitate personalized management strategies. High-risk patients identified at diagnosis could receive closer monitoring, earlier consideration for surgery, and tailored counseling, potentially reducing the incidence of rupture and associated morbidity.

Strengths of this study include its relatively large sample size for a single tertiary center, simultaneous evaluation of multiple inflammatory indices, and the methodological rigor in comparing classical statistical methods with diverse ML algorithms. The use of two independent feature importance approaches (mean decrease in impurity and permutation) strengthens the validity of our predictive variable ranking.

Study Limitations

However, several limitations should be noted. The retrospective design carries an inherent risk of selection bias. All data were derived from a single center, which may limit generalizability. Potential confounders, such as subclinical infections or inflammatory comorbidities, could influence hematologic indices but were not systematically excluded. Although model performance was high, external validation in multicenter prospective cohorts is necessary before clinical adoption.

Future research should focus on validating these findings in larger and more diverse populations, integrating additional biomarkers (e.g., cytokines, cell-free DNA), and developing real-time decision support tools within electronic medical record systems. Randomized controlled trials assessing whether ML-guided treatment decisions improve clinical outcomes compared to current practice would be of particular value. Additional limitations include the single-center design, the potential residual confounding affecting inflammatory markers, and the lack of external validation. Therefore, prospective multicenter studies are required before routine clinical implementation.

Conclusion

This study indicates that AISI, SIRI, and FAR are strong predictors of MTX outcomes in tubal ectopic pregnancy. Integrating these biomarkers into machine-learning models, particularly Naïve Bayes and logistic regression, significantly enhances predictive accuracy and supports individualized risk stratification. The routine use of these indices in clinical practice may facilitate earlier decision-making, optimize treatment selection, and reduce complication rates. However, multicenter prospective studies are required to validate these findings and to explore the added value of advanced imaging and novel molecular biomarkers.

Ethics

Ethics Committee Approval: The study protocol was reviewed and approved by the Ankara Bilkent City Hospital Institutional Ethics Committee (approval no: TABED 2-25-1311, date: 11.06.2025), and all procedures were conducted in accordance with the Declaration of Helsinki.

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: U.Z., G.K., M.İ.H., B.E., Concept: U.Z., Design: U.Z., Data Collection or Processing: G.K., M.İ.H., Analysis or Interpretation: U.Z., B.E., Literature Search: U.Z., S.A.D., G.K., M.İ.H., B.E., Writing: U.Z., S.A.D., G.K., M.İ.H., B.E.

Conflict of Interest: No conflict of interest was declared by the authors.

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

Zorlu et al. Inflammation and ML in ectopic pregnancy

Turk J Obstet Gynecol 2026;23(1):47-55

References

1. Varma R, Gupta J. Tubal ectopic pregnancy. BMJ Clin Evid. 2012;2012:1406.

2. Gervaise A, Capella-Allouc S, Audibert F, Rongières-Bertrand C, Vincent Y, Fernandez H. Methotrexate for the treatment of unruptured tubal pregnancy: a prospective nonrandomized study. JSLS. 2003;7:233-8.

3. Baskiran Y, Uçkan K, Karacor T, Çeleğen I, Acar Z. The impact of maternal electrolyte and albumin levels on the efficacy of single-dose methotrexate treatment for ectopic pregnancies. Turk J Obstet Gynecol. 2023;20:214-8.

4. Tong S, Skubisz MM, Horne AW. Molecular diagnostics and therapeutics for ectopic pregnancy. Mol Hum Reprod. 2015;21:126-35.

5. Dinc K, Issın G. Novel marker to predict rupture risk in tubal ectopic pregnancies: the systemic immune-inflammation index. Ginekol Pol. 2023;94:320-5.

6. Akkaya H, Uysal G. Can hematologic parameters predict treatment of ectopic pregnancy? Pak J Med Sci. 2017;33:937-42.

7. Baskiran Y, Uçkan K, Çeleğen I. Can failure be predicted in methotrexate treatment with the modified parameter? Arch Gynecol Obstet. 2024;310:477-83.

8. Bilir C, Soysal C, Biyik I, Ulas O, Erbakirci NM, Sari H, et al. The relationship between hemogram-based inflammatory indices and prognosis in ectopic pregnancy cases treated with methotrexate. Sci Rep. 2025;15:23114.

9. Seyfettinoglu S, Adiguzel FI. Prediction of tubal rupture in ectopic pregnancy using methotrexate treatment protocols and hematological markers. J Clin Med. 2023;12:6459.

10. Dereli ML, Savran Üçok B, Ozkan S, Sucu S, Topkara S, Firatligil FB, et al. The importance of blood-count-derived inflammatory markers in predicting methotrexate success in patients with tubal ectopic pregnancy. Int J Gynecol Obstet. 2024;167:789-96.

11. Chen S, Chen XF, Qiu P, Huang YX, Deng GP, Gao J. Association between white blood cells at baseline and treatment failure of methotrexate for ectopic pregnancy. Front Med (Lausanne). 2021;8:722963.