Clin. Lab. 2026;72:XXX-XXX

ORIGINAL ARTICLE

The Application of Machine Learning in the Formation Mechanism of Eosinophilic Pleural Effusion

Xu Yuan, Ming Huang, Hui Zhang, Guo Li, Hongbo Zhang

Department of Laboratory Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, China

ABSTRACT

Background: **The aim of this study was to establish a critical point model for the transudation of eosinophils through the pleura using machine learning techniques. This model treats the presence or absence of eosinophils in pleural effusion as a binary variable and incorporated relevant detection indicators to enhance the understanding of the formation mechanism of eosinophilic pleural effusion.**

Methods: **We utilized nearly 50 indicators from 7 major detection items and compared various machine learning models, including XGBoost, Logistic Regression, LightGBM, Random Forest, AdaBoost, Decision Tree, GBDT, and GNB, to identify the optimal model based on appropriate indicators.**

Results: **The AUC (95% CI) values for the validation sets of the XGBoost, Logistic Regression, LightGBM, Random Forest, AdaBoost, Decision Tree, GBDT, and GNB models were 0.714 (0.658 - 0.771), 0.704 (0.647 - 0.761), 0.732 (0.677 - 0.787), 0.743 (0.689 - 0.798), 0.747 (0.694 - 0.801), 0.628 (0.577 - 0.678), 0.767 (0.717 - 0.818), and 0.693 (0.637 - 0.748), respectively. Through the comparison of multiple models, the GBDT model was optimized to include the following seven indicators: APTT, TT, INR, PFTP, PTLDH, PBHGB, and DBIL. The AUC (95% CI) values for the test set were 0.902 (0.885 - 0.920), 0.761 (0.678 - 0.843), and 0.759 (0.708 - 0.811), indicating good generalization.**

Conclusions: **The GBDT model, identified by the code 4LE174761WZ754ba5qT1D, demonstrates that under the appropriate algorithm with the selected macroscopic indicators (APTT, TT, INR, PFTP, PTLDH, PBHGB, and DBIL), the microscopic environment is conducive to the transudation of eosinophils through the pleura, thereby optimizing the understanding of the formation mechanism of eosinophilic pleural effusion.**

(Clin. Lab. 2026;72:xx-xx. DOI: 10.7754/Clin.Lab.2025.250679)

Correspondence:
Hongbo Zhang
Department of Laboratory Medicine
Tongji Hospital
Tongji Medical College
Huazhong University of Science and Technology
Wuhan, 430030
Hubei
China
Phone: + 86 2783662915
Email: 58639096@qq.com
ORCID: 0000-0002-5543-6632

Manuscript accepted August 13, 2025

KEYWORDS

machine learning, GBDT model, eosinophilic pleural effusion, pleural effusion, eosinophil

INTRODUCTION

Eosinophilic pleural effusion (EPE), characterized by the presence of eosinophils constituting ≥ 10% of the pleural effusion [1-3], represents a relatively rare clinicopathological condition that has attracted significant attention within the medical community. The pathogenesis of EPE involves the transmigration of eosinophils across the pleural membrane. However, the extent of pleural injury required to facilitate this cellular move-

ment remains unclear, as do the indicators that can effectively monitor this critical process. Utilizing a decade of historical data and employing advanced machine learning techniques, this study sought to identify macroscopic indicators that reflect the microenvironmental conditions favorable for eosinophilic migration into the pleural cavity, thereby elucidating the mechanisms responsible for the development of eosinophilic pleural effusion.

MATERIALS AND METHODS

Patient filter criteria

1) Data spanning from 2013 through 2024 were selected from the hospital's medical records database, Yidu Cloud.
2) Cases were identified in pleural effusion routine tests where eosinophil percentages were zero or higher.
3) The patient's diagnosis was based on the discharge summary, which provides a comprehensive diagnosis by physicians who integrate diagnostic criteria for multiple diseases with all inpatient test findings. This diagnosis is regarded as highly reliable due to its rigorous methodology and evidence-based foundation. In instances of multiple etiologies, classification was determined by the primary disease that was treated.
4) The collected data encompassed the patient's blood routine, blood biochemistry, coagulation routine, pleural effusion routine, pleural effusion biochemistry, tumor markers, and inflammatory markers. The results were synchronized with the time axis of the pleural effusion routine, selecting those within a three-day range. To minimize the treatment's impact on outcomes, results prior to the standard pleural effusion time point were prioritized.

Selection of detection indicators

A total of seven categories of tests were collected and organized as follows: Pleural fluid route (comprising 10 sub-items), biochemistry of pleural fluid (comprising 6 sub-items), blood routine (comprising 9 sub-items), coagulation routine (comprising 7 sub-items), biochemical examination (comprising 12 sub-items), tumor markers (comprising 7 sub-items), and inflammatory markers (comprising 4 sub-items). Out of these, seven items were found to be closely related to the modeling process. Specifically, within the coagulation routine, three items were identified: activated partial thromboplastin time (APTT), prothrombin time (PT), and international normalized ratio (INR). Within the biochemistry of pleural fluid, two items were pertinent: Pleural fluid total protein (PFTP) and pleural fluid lactate dehydrogenase (PFLDH). the blood routine and the biochemical examination each contributed one item, namely hemoglobin (PBHGB) and direct bilirubin (DBIL), respectively. Although other sub-items were initially considered in the modeling process, they were ultimately excluded and are not enumerated further in this context.

Selection and evaluation of machine learning models

There are various platforms for machine learning, and the types of models available on each platform also vary [4-8]. In this case, the machine learning-related module from the Beckman platform (https://www.xsmartanalysis.com/beckman/login/) was selected for the research. The dataset was randomly partitioned into two subsets: 80% for training and 20% for testing. The training subset underwent tenfold cross-validation, wherein the entire dataset was divided into ten folds. In each iteration, nine folds were utilized as the training set to develop the model, while the remaining fold served as the internal validation set for model evaluation. This process was repeated ten times. We conducted a comparative analysis of early classification models constructed using various machine learning algorithms, including Extreme Gradient Boosting (XGBoost), Logistic Regression, Light Gradient Boosting Machine (LightGBM), Random Forest, Adaptive Boosting (AdaBoost), Decision Tree, Gradient Boosted Decision Trees (GBDT), and Gaussian Naive Bayes (GNB). The objective was to identify the most effective algorithms and early classification models based on their performance on the validation set. The detailed fitting process is illustrated in the schematic diagram (Figure 1).

Shapley Additive Explanations (SHAP) analysis was employed to elucidate the significance and contributions of individual features within the model. To assess the model's performance, accuracy was calculated as a primary metric. The interpretability of the model was further evaluated using several key indicators: 1) Area under the receiver operating characteristic curve (AUC): This metric measures the area under the ROC curve, with values ranging from 0 to 1, indicating the model's ability to distinguish between classes. 2) Cutoff: This refers to the threshold value utilized by the model to differentiate between positive (e.g. eosinophils pass through the pleura) and negative (e.g. eosinophils cannot pass through the pleura) cases. 3) Accuracy: Defined as the proportion of correctly predicted samples, encompassing both positive and negative cases. 4) Sensitivity (Recall): This metric assesses the model's capability to accurately identify positive cases. 5) Specificity: This indicator measures the model's ability to correctly identify negative cases. 6) Positive predictive value (Precision, PPV): This represents the proportion of predicted positive cases that are true positives. 7) Negative predictive value (NPV): This denotes the proportion of predicted negative cases that are true negatives. 8) F1 score: Calculated as the harmonic mean of sensitivity and precision, the F1 score is particularly useful for evaluating models on imbalanced datasets, using the formula F1 = 2 * Precision * Recall/(Precision + Recall). These indicators collectively provide a comprehensive evaluation of the model's interpretability and performance.

In practical modeling applications, the range of AUC values is a crucial metric for assessing a model's generalization capability. Specifically, an AUC value between 0.5 and 0.7 suggests weak generalization performance, while a range of 0.7 to 0.85 indicates relatively strong generalization ability. A value between 0.85 and 0.95 reflects excellent generalization capacity. However, an AUC value between 0.95 and 1 is typically considered unrealistic in real-world scenarios.

Statistical analysis

Continuous variables were reported as either mean ± standard deviation (SD) or median with interquartile range (IQR), contingent upon the data distribution. The Shapiro-Wilk test was employed to assess the normality of the data, which determined the decision to use parametric tests. For two-group comparisons of unpaired nonparametric data, the Mann-Whitney U test was utilized, while the Kruskal-Wallis test was applied for nonparametric comparisons involving more than two groups. Correlation analysis was conducted in conjunction with normality assessments to determine the appropriate use of either the Spearman or Pearson correlation coefficients. Categorical variables were analyzed using the chi-squared (χ^2) test, depending on sample size and expected cell frequencies. Receiver operating characteristic (ROC) curve analysis was performed to determine the optimal cutoff values for parameters, with the objective of maximizing sensitivity and specificity. Modeling was executed using various algorithms, including XGBoost, logistic regression, LightGBM, Random Forest, AdaBoost, Decision Tree, Gradient Boosting Decision Trees (GBDT), and Gaussian Naive Bayes (GNB). The optimal model was subsequently identified. Shapley Additive Explanations (SHAP) values were utilized to elucidate the contribution of each feature within the model. Statistical analyses were performed using GraphPad Prism V.10.0 (San Diego, California, USA). Statistical significance was determined as $p < 0.05$.

RESULTS

Baseline analysis of modeling indicators

In this study, a dataset comprising 2,769 cases was collected, with each case encompassing 55 detection indicators categorized into seven major groups. These indicators were classified into two groups based on the presence or absence of eosinophils, and statistical analyses were conducted to compare the significant differences between the two groups for each indicator. Given the substantial volume of data, the seven primary indicators selected for modeling are presented in the Table 1.

Multi-model comparison

The presence or absence of eosinophils was treated as a binary variable for feature value screening of the indicators. Predictive models were developed using several algorithms, including XGBoost, Logistic Regression, LightGBM, Random Forest, AdaBoost, Decision Tree, Gradient Boosting Decision Tree (GBDT), and Gaussian Naive Bayes (GNB). The primary metric for model evaluation was the AUC, complemented by assessments of accuracy, sensitivity, specificity, PPV, NPV, and the F1 score. The optimal model was subsequently identified. A comparative analysis of the model performance metrics is presented in Table 2.

Summary of parameters related to the test set, validation set, and training set of the optimal GBDT model

Following a comprehensive comparison of various models, the Gradient Boosted Decision Tree (GBDT) model was identified as the most optimal. The model incorporated the presence or absence of eosinophils as a key variable and utilized a combination of 55 detection indicators. From these, seven specific indicators - APTT, TT, INR, PF-TP, PF-LDH, PB-HGB, and DBIL - were selected for the modeling process. The parameters pertaining to the training set, validation set, and test set utilized during the modeling process are detailed in Table 3.

Explanation of the GBDT model

Following the establishment of the GBDT model, its reliability was elucidated through an analysis incorporating five dimensions: Decision Curve, SHapley Additive Explanations (SHAP) values, KS Statistic Plot, SHAP Feature Importance Plot, and SHAP Single Sample Prediction. A comprehensive analysis of the results is conducted in the text below using five pictures.

In the process of evaluating the model, decision curves are often used [5]. Figure 2 shows the decision curve for the GBDT model: The x-axis represents the decision threshold probability, ranging from 0% to 100%, while the y-axis denotes the mean net benefit, which accounts for both gains and costs. The "Treat None" line illustrates the net benefit when no intervention is implemented, serving as the baseline. Conversely, the "Treat All" line indicates the net benefit when interventions are applied universally across all samples. The "GBDT" line reflects the net benefit derived from employing this model to inform decision-making. Analysis of the figure reveals that when the threshold probability ranges from approximately 20% to 80%, the GBDT model yields a higher net benefit compared to the "Treat None" and "Treat All" strategies, reaching a maximum of approximately 0.23. Additionally, at around a 90% threshold probability, the model demonstrates a notable advantage over the baseline, as evidenced by the small sharp peaks in the figure.

SHAP value is an important metric of the model, reflecting the role of the metric in the model [5,9]. Figure 3 shows the summary graph of SHAP (SHapley Additive Explanations) analysis: APTT emerges as the most significant predictor, whereas liver function indicators, specifically DBIL, exhibit comparatively minor

Table 1. Baseline characteristics of model indicators.

Item	Total (n = 2,769)	PF-EOS% = 0 (n = 950)	PF-EOS% ≥ 1 (n = 1,819)	p
APTT, median [IQR]	38.400 [34.400, 42.900]	35.200 [31.200, 41.000]	39.400 [36.100, 43.500]	< 0.001
TT, median [IQR]	16.100 [15.200, 17.100]	15.500 [14.200, 16.800]	16.300 [15.500, 17.100]	< 0.001
INR, median [IQR]	1.100 [1.020, 1.200]	1.150 [1.060, 1.250]	1.080 [1.010, 1.170]	< 0.001
DBIL, median [IQR]	3.400 [2.300, 5.600]	4.000 [2.600, 6.700]	3.200 [2.200, 5.100]	< 0.001
PFLDH, median [IQR]	259.000 [160.000, 454.000]	229.000 [125.000, 433.000]	279.000 [176.000, 462.000]	< 0.001
PBHGB, median [IQR]	119.000 [102.000, 134.000]	113.000 [94.000, 128.000]	123.000 [106.000, 136.000]	< 0.001
PFTP, median [IQR]	43.500 [33.800, 50.300]	39.700 [27.000, 48.500]	44.600 [37.600, 50.800]	< 0.001

Table 2. Data comparison of multi-model validation sets.

Model types	AUC (95% CI)	Cutoff (95% CI)	Accuracy (95% CI)	Sensibility (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	F1 score (95% CI)
XGBoost	0.714 (0.658 - 0.771)	0.732 (0.709 - 0.755)	0.701 (0.691 - 0.712)	0.758 (0.734 - 0.782)	0.576 (0.545 - 0.606)	0.796 (0.788 - 0.804)	0.524 (0.509 - 0.539)	0.776 (0.763 - 0.789)
Logistic	0.704 (0.647 - 0.761)	0.665 (0.656 - 0.674)	0.7 (0.682 - 0.718)	0.753 (0.732 - 0.775)	0.58 (0.556 - 0.604)	0.798 (0.781 - 0.816)	0.516 (0.492 - 0.541)	0.775 (0.758 - 0.792)
LightGBM	0.732 (0.677 - 0.787)	0.614 (0.606 - 0.622)	0.737 (0.728 - 0.745)	0.825 (0.815 - 0.835)	0.542 (0.518 - 0.567)	0.799 (0.789 - 0.809)	0.584 (0.563 - 0.606)	0.812 (0.806 - 0.817)
Random Forest	0.743 (0.689 - 0.798)	0.575 (0.559 - 0.591)	0.749 (0.735 - 0.762)	0.852 (0.830 - 0.873)	0.53 (0.503 - 0.556)	0.795 (0.782 - 0.807)	0.629 (0.596 - 0.662)	0.822 (0.810 - 0.833)
AdaBoost	0.747 (0.694 - 0.801)	0.503 (0.503 - 0.504)	0.724 (0.709 - 0.739)	0.782 (0.757 - 0.808)	0.603 (0.574 - 0.632)	0.805 (0.788 - 0.822)	0.571 (0.545 - 0.598)	0.793 (0.778 - 0.807)
Decision Tree	0.628 (0.577 - 0.678)	1.0 (1.000 - 1.000)	0.668 (0.656 - 0.680)	0.741 (0.726 - 0.756)	0.514 (0.484 - 0.544)	0.764 (0.748 - 0.780)	0.483 (0.465 - 0.501)	0.752 (0.742 - 0.762)
GBDT	0.767 (0.717 - 0.818)	0.657 (0.643 - 0.671)	0.736 (0.723 - 0.749)	0.795 (0.769 - 0.822)	0.615 (0.581 - 0.649)	0.808 (0.792 - 0.824)	0.6 (0.574 - 0.626)	0.801 (0.787 - 0.814)
GNB	0.693 (0.637 - 0.748)	0.938 (0.925 - 0.951)	0.643 (0.630 - 0.655)	0.63 (0.606 - 0.655)	0.667 (0.638 - 0.695)	0.804 (0.792 - 0.816)	0.455 (0.437 - 0.473)	0.706 (0.690 - 0.723)

Table 3. Parameters for the training, validation and test set of the GBDT model.

Item	AUC (95% CI)	Cutoff (95% CI)	Accuracy (95% CI)	Sensibility (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	F1 score (95% CI)
Training	0.902 (0.885 - 0.920)	0.653 (0.638 - 0.669)	0.857 (0.850 - 0.864)	0.900 (0.887 - 0.914)	0.762 (0.749 - 0.775)	0.893 (0.889 - 0.897)	0.778 (0.756 - 0.801)	0.897 (0.891 - 0.902)
Validation	0.761 (0.678 - 0.843)	0.653 (0.638 - 0.669)	0.748 (0.720 - 0.775)	0.814 (0.787 - 0.842)	0.601 (0.549 - 0.652)	0.818 (0.797 - 0.840)	0.597 (0.551 - 0.642)	0.816 (0.795 - 0.837)
Test	0.759 (0.708 - 0.811)	0.666	0.729	0.789	0.617	0.795	0.608	0.792

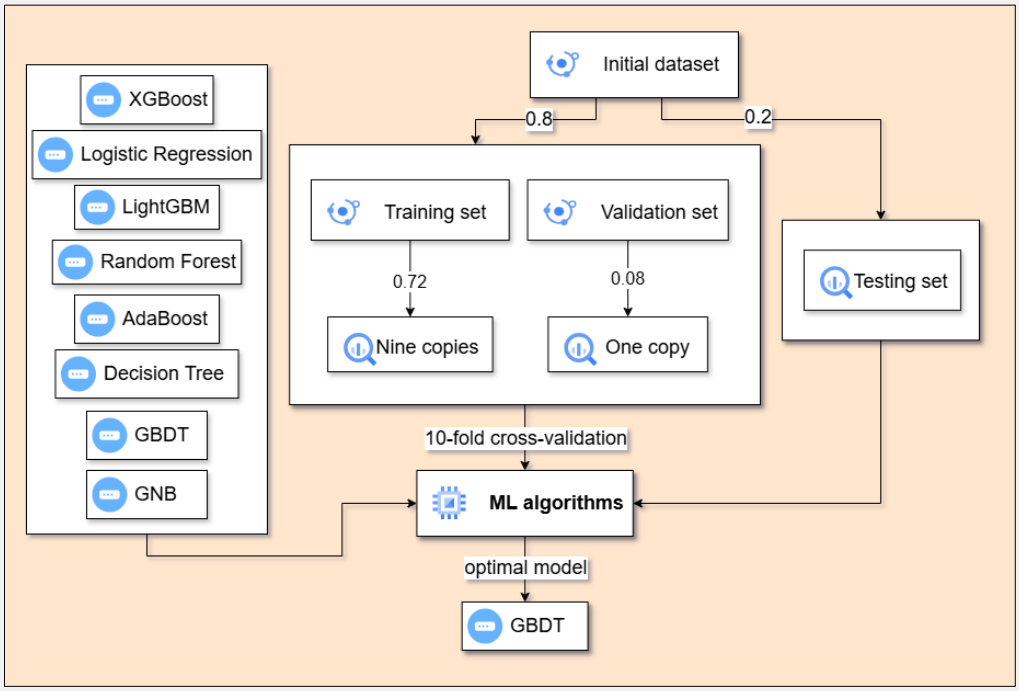

Figure 1. Flow chart for comparing and constructing machine learning models.

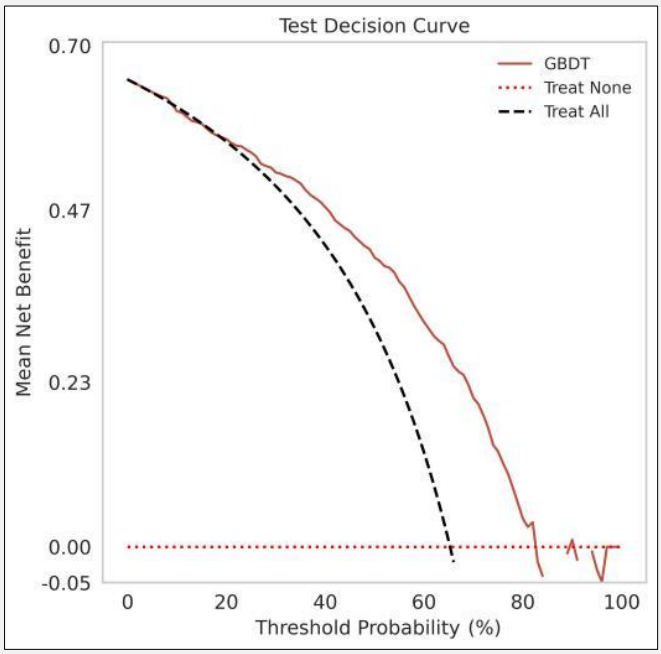

Figure 2. Decision curve of the GBDT model.

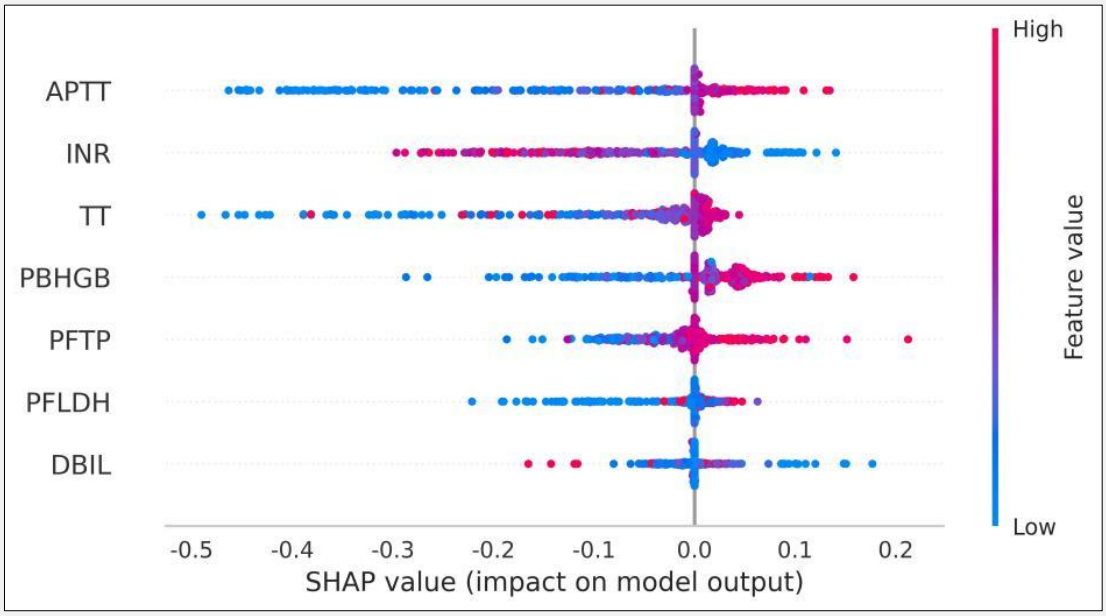

Figure 3. Summary graph of SHapley Additive Explanations analysis.

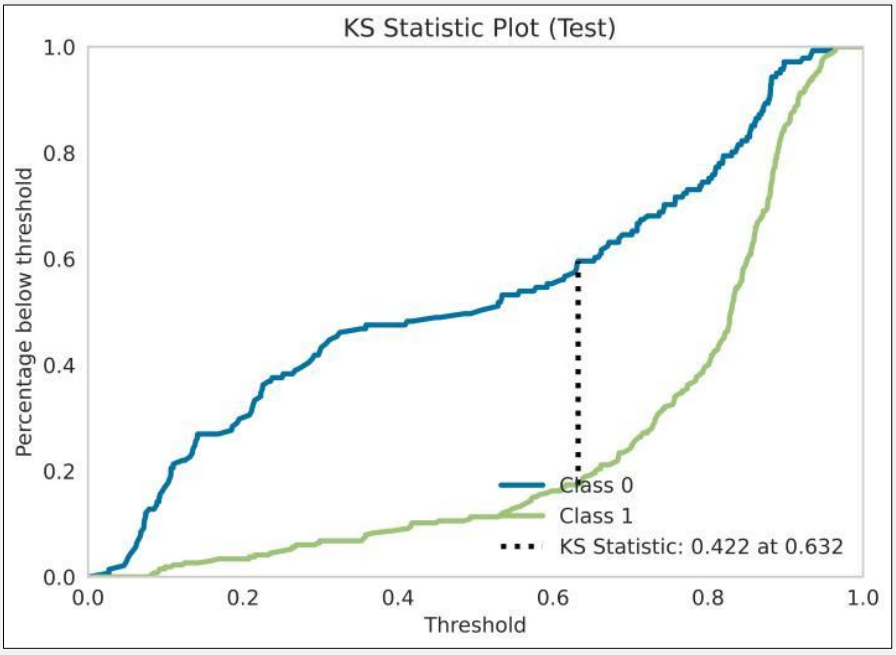

Figure 4. KS statistic plot explanation.

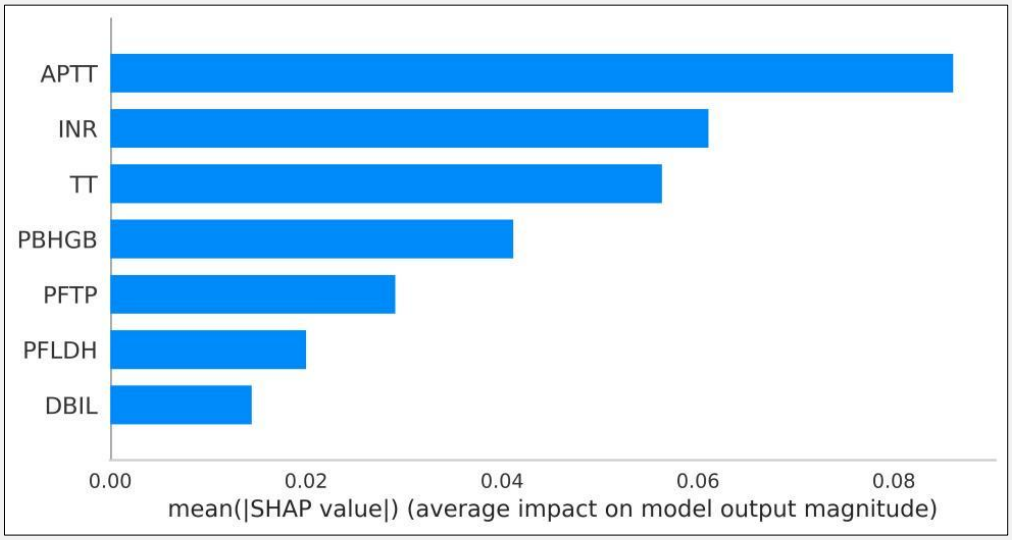

Figure 5. SHAP feature importance plot explanation.

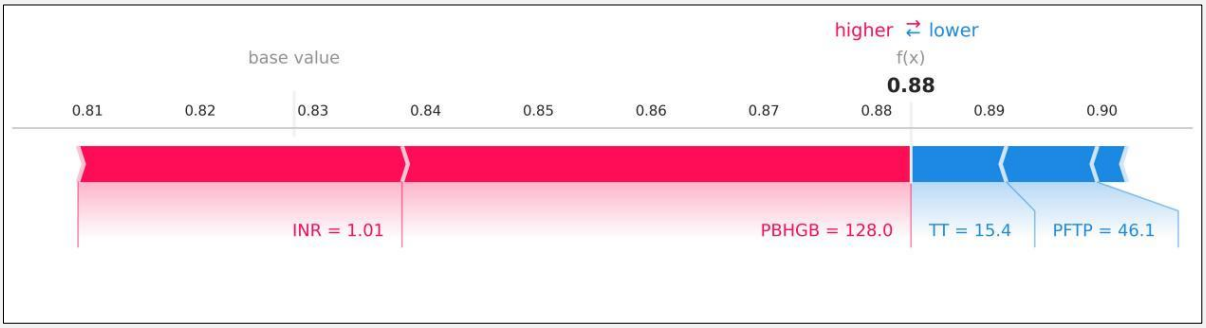

Figure 6. SHAP single sample prediction explanation chart.

influence. Lower values, correspond to negative SHAP values, while higher values, correspond to positive SHAP values. The SHAP values for INR and DBIL are predominantly negative, suggesting that an increase in these indicator values is likely to decrease the model's predicted output. Conversely, the SHAP values for APTT, TT, PBHGB, PFTP, and PFLDH are primarily positive, indicating that an increase in these indicator values is likely to enhance the model's predicted output.

The discrimination of the model for the studied events requires the use of KS statistics. Figure 4 shows the KS Statistic Plot (Kolmogorov-Smirnov Plot) explanation of the GBDT model: This graph illustrates the outcomes of the Kolmogorov-Smirnov (KS) test, employed to evaluate the discriminatory capacity of predicted scores from a binary classification model (distinguishing between Class 0 and Class 1). The optimal cutoff point is identified at a threshold of 0.632, where the KS statistic attains its maximum value of 0.422. This threshold represents the most effective balance for utilizing the model in classification tasks. Generally, a KS value exceeding 0.3 is indicative of a model's ability to discriminate between classes. A KS value of 0.422 implies that the model can reasonably differentiate between the two

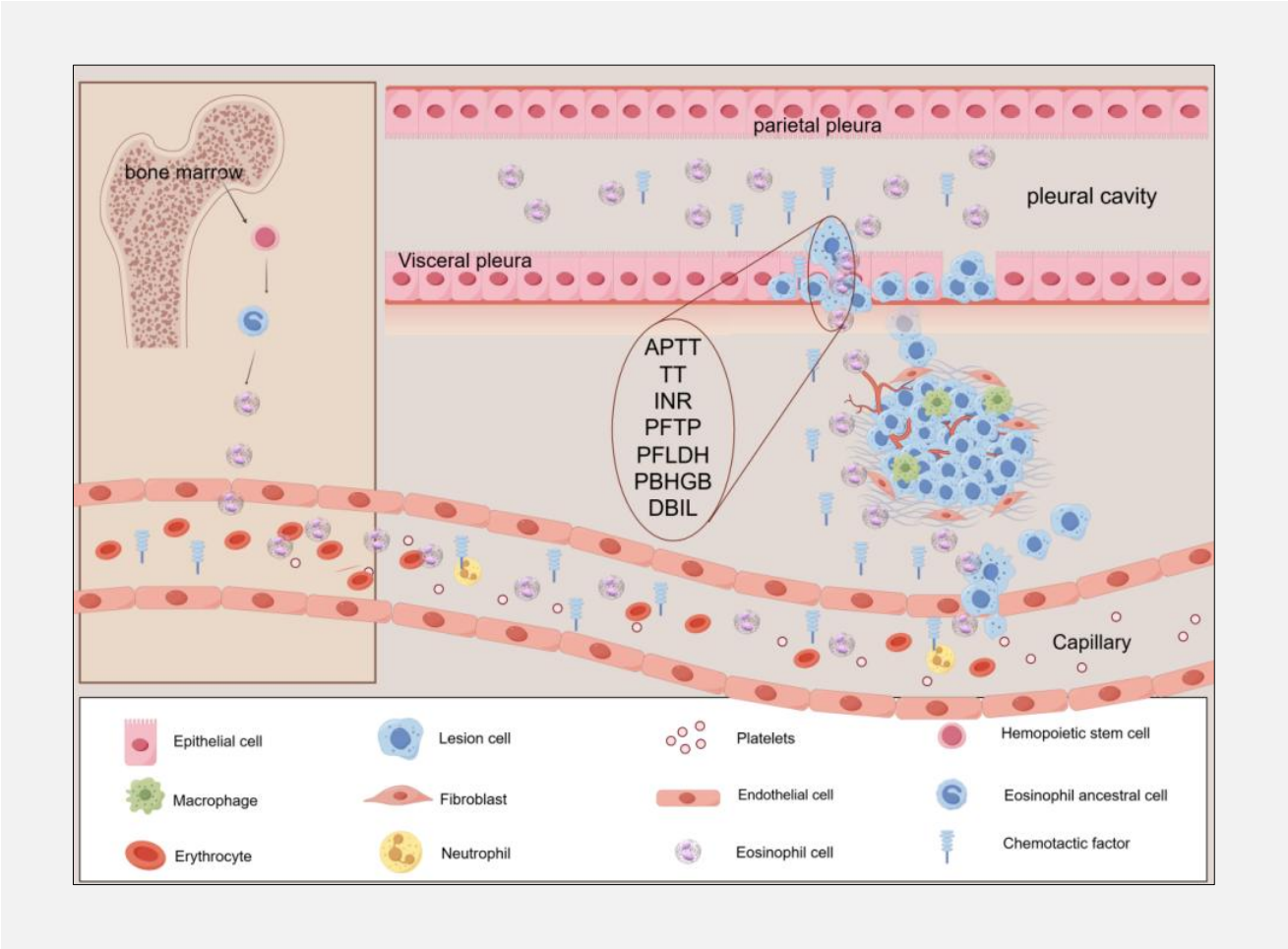

Figure 7. Mechanism of eosinophilic pleural effusion formation.

sample types, although it does not achieve an exceptional level of performance, as outstanding models typically exhibit KS values greater than 0.6.

The contribution degrees of each index for constructing the model to the model are different. At this time, the average impact factor in the SHAP value needs to be used for explanation [8]. Figure 5 shows the SHAP Feature Importance Plot Explanation: This figure presents a feature importance ranking chart derived from SHAP values, illustrating the mean impact magnitude of various blood test indicators on the model's predictive outcomes. It complements the data provided by the preceding SHAP summary plot by quantifying the average impact magnitude of each feature. The chart reaffirms the central role of APTT, consistent with the findings of the summary plot. However, it should be noted that this chart solely represents the intensity of the impact and does not convey the direction (positive or negative) of the influence.

The original intention of establishing a model usually is to find the best decision point [9], that is, a corresponding single-sample prediction point in the model. Figure 6 shows the SHAP Single Sample Prediction Explanation Chart output by the GBDT model: This chart illustrates the SHAP analysis, elucidating the prediction for an individual sample (patient) by visually demonstrating the collective contribution of each feature to the final model prediction. The predicted value for this sample, ranging from 0.88 to 0.90, exceeds the baseline value, which lies between 0.81 and 0.90. Specifically, the features INR = 1.01 and PBHGB = 128 positively influence the prediction, whereas TT = 15.4 seconds and PFTP = 46.1 contribute negatively. This interplay of features results in a balanced prediction, with the probability of eosinophils passing through the pleura being 0.88, or 88%.

External validation of the GBDT model

Under typical circumstances, the completion of model design necessitates external validation, as the model's logic is predicated on the need to evaluate future data. However, this study deviated slightly from the conventional approach. The primary objective here was to identify the conditions under which eosinophils can

penetrate the pleura, eliminating the requirement for future predictions. Consequently, the study utilized the available data solely for modeling purposes, employing the model's test set to assess its effectiveness. Therefore, external validation was not conducted, and all data were utilized for modeling. This approach represents a slight difference from the traditional modeling process.

DISCUSSION

The primary objective of this article was to ascertain the conditions under which eosinophils infiltrate the pleura and subsequently lead to the development of eosinophilic pleural effusion, utilizing a modeling approach. A critical aspect of this investigation involved the systematic organization and analysis of the data. In this study, a comprehensive dataset comprising 2,769 data points was collected, categorized into 7 major groups and 55 detection indicators.

To facilitate a more nuanced understanding, the dataset was bifurcated based on the presence or absence of eosinophils in the pleura. Given the extensive nature of the 55 indicators, the study concentrated on 7 key indicators for subsequent modeling: APTT, TT, INR, PFTP, PFLDH, PBHGB, and DBIL, as detailed in Table 1. The analysis revealed that the p-values for both groups were less than 0.001, signifying that these 7 indicators exhibit highly significant differences between the two groups. This finding indicates the high sensitivity of these indicators in discerning the penetration of eosinophils into the pleura.

Following a comprehensive analysis and organization of the data, it was imperative to select appropriate indicators for constructing various models. The evaluation of these models was conducted using the area under the curve (AUC) as the primary metric, supplemented by additional metrics such as cutoff, accuracy, sensitivity (Recall), specificity, positive predictive value (PPV or Precision), negative predictive value (NPV), and the F1 score. As illustrated in Table 2, the AUC values for the eight models - XGBoost, Logistic Regression, LightGBM, Random Forest, AdaBoost, Decision Tree, Gradient Boosting Decision Tree (GBDT), and Gaussian Naive Bayes (GNB) - within the 95% confidence interval were as follows: 0.714 (0.658 - 0.771), 0.704 (0.647 - 0.761), 0.732 (0.677 - 0.787), 0.743 (0.689 - 0.798), 0.747 (0.694 - 0.801), 0.628 (0.577 - 0.678), 0.767 (0.717 - 0.818), and 0.693 (0.637 - 0.748), respectively. Notably, the GBDT model exhibited the highest AUC value of 0.767 (0.717 - 0.818). Furthermore, it demonstrated superior PPV and F1 score values, measured at 0.808 (0.792 - 0.824) and 0.801 (0.787 - 0.814), respectively. Consequently, employing the GBDT model in conjunction with seven specific indicators - APTT, TT, INR, PFTP, PFLDH, PBHGB, and DBIL - yielded the most effective results. This model optimally reflects the critical state of eosinophil translocation across the pleura.

With the establishment of the GBDT model, we moved on to the "binary classification model" module of machine learning. Cross-validation was employed as the validation technique, utilizing a random seed of 42, with 10 iterations (folds) and a test set ratio of 0.2. The classification was performed using the GBDT method, and the model's performance was assessed using the receiver operating characteristic - area under the curve (ROC-AUC) metric. Evaluation indicators included APTT, TT, INR, PFTP, PFLDH, PBHGB, and DBIL. The specific parameters of the GBDT model are detailed in Table 3. The performance metrics for the training, validation, and test sets were 0.902 (95% CI: 0.885 - 0.920), 0.761 (95% CI: 0.678 - 0.843), and 0.759 (95% CI: 0.708 - 0.811), respectively. According to the model generalization standards, a range of 0.5 - 0.7 indicates weak generalization, 0.7 - 0.85 suggests relatively strong generalization, 0.85 - 0.95 signifies excellent generalization, and values between 0.95 - 1 are typically unrealistic in practical applications. The test set performance of 0.759 for the GBDT model indicates a strong generalization capability.

Furthermore, it is imperative to demonstrate the model's reliability through a comprehensive analysis of detailed parameters, as shown in Figure 2 to Figure 6. Figure 2, the decision curve, illustrates that the model offers substantial clinical decision-making value within the intermediate probability threshold range of 20% to 80%. In Figure 3, the summary graph of SHAP (SHapley Additive Explanations) analysis reveals that the SHAP values for INR and DBIL were predominantly negative, implying that an increase in these indicators is associated with a reduction in the model's predicted output. Conversely, the SHAP values for APTT, TT, PBHGB, PFTP, and PFLDH were primarily positive, indicating that an increase in these indicators is likely to enhance the model's predicted output. Consequently, to mitigate the incidence of eosinophilic pleural effusion, it is advisable to manage and limit the elevation of APTT, TT, PBHGB, PFTP, and PFLDH indicators. In the context of Figure 4, the KS statistic plot (Kolmogorov-Smirnov Plot) demonstrates that the model achieved a maximum KS statistic value of 0.422 at the optimal cutoff point of 0.632. Although it is still short of the excellent model's KS > 0.6, 0.422 indicates that the model has a good differentiating ability. In the analysis of Figure 5 (SHAP feature importance chart), this visualization effectively illustrates the significance of each indicator within the model [8]. Specifically, APTT occupies a central role, exerting the most substantial influence. This suggests that APTT is a critical indicator of pleural damage. A plausible explanation for this is that damage to mesothelial cells in the pleura may result in the secretion of cytokines, thereby causing abnormal APTT levels. Figure 6 offers a SHAP-based explanation for a single sample prediction, providing an illustrative example of the model's functioning. The final predicted value for this sample (ranging from 0.88 to 0.90) slightly exceeds the baseline value (0.81 to 0.90). In this context, an in-

ternational normalized ratio (INR) of 1.01 and a peripheral blood hemoglobin (PBHGB) level of 128 contributed positively to the prediction, whereas a thrombin time (TT) of 15.4 seconds and a plasma fibrinogen time (PFTP) of 46.1 contributed negatively, achieving an equilibrium. At this equilibrium, the probability of the event occurring was 0.88, indicating an 88% likelihood of eosinophils traversing the pleura. The detailed interpretation of each parameter underscores the GBDT model's robust generalization capabilities and reliability.

With the acquisition of the GBDT model, whose code is 4LE174761WZ754ba5qT1D, we now resume our investigation into eosinophilic pleural effusion. The preliminary modeling efforts were also directed towards this objective. Eosinophilic pleural effusion can occur when the pleura is stimulated by factors such as trauma (e.g., pneumothorax, hemothorax, or surgery), infections (e.g., parasitic infections, tuberculosis), malignant tumors (e.g., lymphoma, metastatic cancer), autoimmune diseases (e.g. eosinophilic granulomatosis with polyangiitis (EGPA)), or adverse drug reactions [1,10–12]. In response to these stimuli, local tissues and immune cells, including Th2 lymphocytes, mast cells, and mesothelial cells, release critical chemokines such as interleukin-5 (IL-5), eotaxin, and granulocyte-macrophage colony-stimulating factor (GM-CSF) [13-15].

These chemokines facilitate the directed migration of eosinophils from the bone marrow and peripheral blood into the pleural cavity. Upon activation within the pleural cavity, eosinophils inflict direct damage to pleural tissue by releasing cytotoxic granular proteins, such as major basic protein (MBP) and eosinophil cationic protein (ECP). Concurrently, they secrete inflammatory mediators, including leukotrienes, platelet-activating factor (PAF), and tumor necrosis factor-alpha (TNF-α), which further enhance vascular permeability and lead to the exudation of plasma components [16]. The increase in vascular permeability establishes conditions conducive to the migration of eosinophils across blood vessel walls and into adjacent tissues. This is accompanied by plasma exudation, which modifies local osmotic pressure and facilitates the development of pleural effusion. However, this process alone allows eosinophils to traverse the vascular wall and remain confined within the connective tissue. The further migration of eosinophils from the connective tissue into the pleural cavity depends on the severity of pleural injury. In this study, the GBDT model, utilizing seven clinical indicators - APTT, TT, INR, PFTP, PFLDH, PBHGB, and DBIL - was applied to determine this critical threshold. At this critical juncture, the disruption of the pleural microenvironment becomes sufficiently pronounced to allow eosinophil translocation across the pleura into the pleural space, thereby contributing to the EPE formation. A visual depiction of this mechanism is presented in Figure 7.

Limitations

This study did not include biomarkers directly linked to pleural injury, such as IL-5, IL-4, IL-13, CCL11, CCL24, and CCL26. Incorporating these indicators into model development could potentially improve its generalizability, as evidenced by a higher AUC in the test set, potentially exceeding the current value of 0.759. Including such indicators in regular diagnostic protocols would greatly benefit research and model performance.

CONCLUSION

By utilizing GBDT modeling, a predictive model was developed to elucidate the capacity of eosinophils to infiltrate the pleura and induce eosinophilic pleural effusion. This was achieved by integrating seven measurable indicators: APTT, TT, INR, PFTP, PFLDH, PBHGB and DBIL.

The model effectively delineates the critical state of eosinophil penetration into the pleura through quantifiable macroscopic indicators, thereby offering a more comprehensive understanding of the pathogenesis of eosinophilic pleural effusion from a macroscopic standpoint. Furthermore, it proposes novel therapeutic approaches for eosinophilic pleural effusion. For example, the model shows that elevated levels of APTT and TT facilitate the development of eosinophilic pleural effusion. However, whether modulating these increases can mitigate the formation of eosinophilic pleural effusion requires empirical validation.

Acknowledgment:

The authors are grateful to the staff of the Laboratory Medicine Department at Tongji Hospital for their essential technical support.

Ethical Approval:

This study was conducted in accordance with the Declaration of Helsinki and relevant national laws and regulations. It received approval from the Medical Ethics Committee of Tongji Hospital, Tongji Medical College (ethics number: TJ-IRB202506088), and adhered to all applicable national regulations and institutional guidelines. Due to its retrospective nature, a waiver of informed consent was implemented.

Data Availability Statement:

The datasets supporting the conclusions of this article are included within the article and tables. Additional data may be available from the corresponding author upon reasonable request.

Source of Funds:

This study did not receive any funding in any form.

Declaration of Interest:

The authors have no affiliations or involvements with any organizations or entities that hold financial interests related to the subject matter or materials presented in this manuscript.

References:

1. Kalomenidis I, Light RW. Eosinophilic pleural effusions. Curr Opin Pulm Med 2003 Jul;9(4):254-60. (PMID: 12806236)

2. Krenke R, Nasilowski J, Korczynski P, et al. Incidence and aetiology of eosinophilic pleural effusion. Eur Respir J 2009 Nov; 34(5):1111-7. (PMID: 19386682)

3. Li M, Zeng Y, Li Y, Jia D, Chen S, Wang J. Incidence, aetiology and clinical features of eosinophilic pleural effusion: a retrospective study. BMC Pulm Med 2021 Dec 6;21(1):402. (PMID: 34872535)

4. Wang Y, Wei W, Ouyang R, et al. Novel multiclass classification machine learning approach for the early-stage classification of systemic autoimmune rheumatic diseases. Lupus Sci Med 2024 Jan 31;11(1):e001125. (PMID: 38302133)

5. Liu Y, Liang Z, Yang J, et al. Diagnostic and comparative performance for the prediction of tuberculous pleural effusion using machine learning algorithms. Int J Med Inform 2024 Feb;182: 105320. (PMID: 38118260)

6. Kim NY, Jang B, Gu KM, Park YS, Kim YG, Cho J. Differential Diagnosis of Pleural Effusion Using Machine Learning. Ann Am Thorac Soc 2024 Feb;21(2):211-7. (PMID: 37788372)

7. Zhou X, Chen Y, Gui W, et al. Enhanced differential evolution algorithm for feature selection in tuberculous pleural effusion clinical characteristics analysis. Artif Intell Med 2024 Jul;153: 102886. (PMID: 38749310)

8. Ren Z, Hu Y, Xu L. Identifying tuberculous pleural effusion using artificial intelligence machine learning algorithms. Respir Res 2019 Oct 16;20(1):220. (PMID: 31619240)

9. Wu C, Liu W, Mei P, et al. The large language model diagnoses tuberculous pleural effusion in pleural effusion patients through clinical feature landscapes. Respir Res 2025 Feb 12;26(1):52. (PMID: 39939874)

10. Li WJ, Lin ZD, Wang JL. A narrative review of malignant eosinophilic pleural effusion: incidence, etiology and prognostic significance. Ann Palliat Med 2021 Feb;10(2):2314-22. (PMID: 33549025)

11. Ferreiro L, San José E, González-Barcala FJ, et al. Eosinophilic pleural effusion: incidence, etiology and prognostic significance. Arch Bronconeumol 2011 Oct;47(10):504-9. (PMID: 21831499)

12. Alagha K, Tummino C, Sofalvi T, Chanez P. Iatrogenic eosinophilic pleural effusion. Eur Respir Rev 2011 Jun;20(120):118-20. (PMID: 21632800)

13. Taniguchi H, Mukae H, Matsumoto N, et al. Elevated IL-5 levels in pleural fluid of patients with paragonimiasis westermani. Clin Exp Immunol 2001 Jan;123(1):94-8. (PMID: 11168004)

14. Kalomenidis I, Stathopoulos GT, Barnette R, et al. Eotaxin-3 and interleukin-5 pleural fluid levels are associated with pleural fluid eosinophilia in post-coronary artery bypass grafting pleural effusions. Chest 2005 Jun;127(6):2094-100. (PMID: 15947325)

15. Takeuchi E, Takahashi N, Morizumi S, et al. Interleukin-5-producing malignant pleural mesothelioma with eosinophilic pleural effusion. Thorac Cancer 2020 Oct;11(10):3043-6. (PMID: 32894005)

16. Wen T, Rothenberg ME. The Regulatory Function of Eosinophils. Microbiol Spectr 2016 Oct;4(5):10.1128/microbiolspec. MCHD-0020-2015. (PMID: 27780017)