

Highlights

Machine learning model for predicting the visit duration of urticaria based on clinical laboratory data

Yijun Yang, Guofang Li, Zhen Zhang, Zhirong Yao, Linting Huang, Hui Zhang

- A machine learning model was built to predict the visit duration of urticaria based on clinical laboratory data.
- Inversed trend of shapley values of laboratory data in different phase of disease was observed.

Machine learning model for predicting the visit duration of urticaria based on clinical laboratory data

Yijun Yang^{a,1}, Guofang Li^{a,1}, Zhen Zhang^a, Zhirong Yao^a, Linting Huang^{b,**}, Hui Zhang^{a,*}

^a*Shanghai XinHua Hospital affiliated to Shanghai Jiao Tong University School of Medicine, Address line of institution one, Shanghai, 200092, Shanghai, China*

^b*The ninth hospital of Shanghai Jiao Tong University School of Medicine, Address line of institution two, Shanghai, 200011, Shanghai, China*

Abstract

Background: Urticaria is a common condition presenting with wheals, angioedema, or both, driven by mast cell degranulation. The duration of urticaria, particularly chronic spontaneous urticaria (CSU), is crucial for effective patient management and treatment planning. This study aimed to build a machine learning model for predicting the visit duration of urticaria based on clinical laboratory data and to identify the factors that contribute to prolonged episodes of chronic urticaria.

Method:

Results:

Conclusion:

Keywords:

1. Introduction

Urticaria is a common condition presenting with wheals, angioedema, or both, driven by mast cell degranulation(Zuberbier et al., 2021; Radonjic-Hoesli et al., 2018; Ring and Grosber, 2012). The lifetime prevalence for acute urticaria is approximately 20% (Zuberbier et al., 2021). Urticaria is classified based on duration and triggers. Acute urticaria lasts less than 6 weeks, often triggered by specific causes like drugs, food, or infections. While chronic urticaria lasts more than 6 weeks and can be further classified into chronic spontaneous urticaria (CSU) and chronic inducible urticaria (CIndU)(Zuberbier et al., 2021; Ring and Grosber, 2012). CSU is characterized by the spontaneous occurrence of wheals and/or angioedema without a specific trigger and often associated with autoimmune mechanisms(Schettini et al., 2023), while CIndU is triggered by specific stimuli like cold, heat, or pressure(Pozderac et al., 2020). The prevalence of CSU is approximately 0.5% in general population, and is less prevalent in children compared to adults(Balp et al., 2015; Poddighe, 2019; Labbene and Tekou, 2023).

Chronic urticaria significantly impairs quality of life, affecting work and school performance. It is considered a severe allergic disease due to its disabling nature and high disease burden(Zuberbier et al., 2021). Duration of CSU greater than 3 years are associated with better responses to second-generation antihistamines and other treatments(Chiang et al., 2022). Predicting the duration of urticaria, particularly chronic spontaneous urticaria (CSU), is crucial for effective patient management and treatment planning.

Several clinical features have been associated with the severity and duration of chronic spontaneous urticaria (CSU). Higher age at onset, female gender, longer disease duration, and hypersensitivity to aspirin or nonsteroidal anti-inflammatory drugs (NSAIDs) are linked to more severe CSU and prolonged time to remission (SanchezBorges et al., 2017; Rabelo-Filardi et al., 2013). Patients exhibiting concomitant CIndU and recurrent angioedema also tend

*Corresponding author. Email: authora@example.com

**Corresponding author. Email: huanglinting@163.com

¹Yijun Yang and Guofang Li contributed equally to this work.

to experience longer durations of CSU (SanchezBorges et al., 2017; Curto-Barredo et al., 2018). Moreover, patients with multiple allergic conditions are more likely to have prolonged episodes of urticaria (Lin et al., 2011).

Potential biomarkers for CSU severity and duration have been identified. Positive autologous serum skin test (ASST) results, basophil counts, levels of inflammatory markers, activation markers of the extrinsic coagulation pathway, immunoglobulin E (IgE), eosinophils counts, and vitamin D levels are all associated with the disease's severity and duration (SanchezBorges et al., 2017; Rabelo-Filardi et al., 2013; Kolkhir et al., 2019). Specifically, plasma levels of prothrombin fragment, D-dimer, and C-reactive protein (CRP) may serve as markers of CSU severity (Rabelo-Filardi et al., 2013). Serum diamine oxidase (DAO) levels have been linked to the response to antihistamines and dietary interventions, indicating a potential role in predicting disease duration (Chiang et al., 2022).

Metabolic factors also play a role, with high waist circumference (WC), rather than high body mass index (BMI), emerging as a predictive risk factor for longer disease duration in CSU patients (Kim et al., 2021).

The aim of this study was to build a machine learning model for predicting the visit duration of urticaria based on clinical laboratory data and to identify the factors can contribute to prolonged episodes of chronic urticaria by analyzing the importance of variables in the model, hoping to provide a reference for the clinical management of urticaria.

2. Methods

2.1. Patients

Patients with urticaria were recruited from the urticaria specialty clinic of the dermatology department of Shanghai XinHua Hospital affiliated to Shanghai Jiao Tong University School of Medicine from January 2018 to December 2024. The inclusion criteria were as follows: (1) patients diagnosed with urticaria according to the EAACI/GA2LEN/EDF/WAO guidelines(Zuberbier et al., 2021); (2) patients with complete clinical and laboratory data; (3) patients with stable follow-up history, indicated by at least 3 times of follow-up visits. the exclusion criteria were as follows: (1) patients with other skin diseases; (2) patients with severe systemic diseases; (3) patients with incomplete clinical data. The study was approved by the ethics committee of Shanghai XinHua Hospital affiliated to Shanghai Jiao Tong University School of Medicine, and all patients provided written informed consent.

2.2. Data collection and management

Data of patients with urticaria were collected from the electronic medical record system of the hospital, including demographic data, clinical data, laboratory data. the data were stored in a Mysql database for subsequent analysis, with schema as follows: (1) Patients: containing basic information of each unique patient; (2) OutpatientNumbers: storing relationship between outpatient numbers with unique patient; (3) PatientVisits: containing visit events records; (4) PatientExaminations: containing the examination events records; (5) ExaminationItems: a dictionary table describing the examination items. the database schema is shown in database. Database markup language (DBML) describing the database schema is shown in supplementary materials Appendix A

2.3. Feature extraction and feature engineering

Visit duration, calculated by the difference between the first visit date and the last visit date in the urticaria specialty clinic, was the target variable. We suppose that the visit duration of a patient at the urticaria specialty clinic is a surrogate for the disease duration of urticaria. The following features were extracted from the database: (1) demographic data: gender, first visit age; (2) clinical data: whether the patient having concomitant inducible urticaria; (3) laboratory data: results from common blood tests, CRP, immunoglobulin, 25-hydroxyvitamin D, Thyroid function, autoantibodies, coagulation function, common urine tests, and allergen specific IgE tests. For laboratory data, 2 types of data were extracted: time-independent data and time-dependent data. Time-independent data are average values of laboratory data during the whole follow-up period, while time-dependent data are average values of laboratory data during preclinical phase (before the onset of urticaria), acute phase (within 6 weeks after the onset of urticaria), and chronic phase (after 6 weeks of the onset of urticaria). 2 datasets were generated: one with time-independent data and one with time-dependent data, and were compared for prediction performance in the model development process. The sql queries for data are shown in in supplementary materials Appendix A and Appendix A.

2.4. Model development and comparison

Dataset was split into training set and external test set with a ratio of 7:3. 5 models were adopted for comparison: Xgboost, random forest, gradient boosting machine (GBM), adaboost, and support vector machine (SVM). Time-dependent data extraction may introduce multicollinearity between features, which can affect model performance. To address this issue, L1 regularization, L2 regularization, and column sampling methods were employed in XGBoost to handle multicollinearity. L1 regularization helps eliminate irrelevant collinear features by sparse feature selection, while L2 regularization reduces the model's dependence on collinear features by smoothing the weights. Column sampling reduces the likelihood of collinear features being selected simultaneously by randomly choosing a subset of features when building each tree, which is helpful when interpreting the model using method like SHAP, since it enables more balanced weight distribution across features and reduces the possibility of one feature overshadowing others. The Random Forest model does not support L1 and L2 regularization, but it handles multicollinearity through the max_features parameter, which is similar to column sampling. L2 regularization was applied to the SVM model as well. Additionally, the frameworks of XGBoost and GBM incorporate greedy split and boosting mechanisms, selecting the most important features for each tree and fitting the residuals step by step, allowing the model to focus on the key features rather than over-relying on collinear ones. Hyperparameter optimization on 10 model settings (5 models and 2 datasets) was performed using TPE algorithm by nni package in python, which is a bayesian optimization algorithm that uses tree-structured parzen estimator to model the objective function and suggest the next set of hyperparameters to evaluate based on the previous results. Internal 5 fold cross-validation was employed to discern the most suitable hyperparameters for each distinct model, individually applied to each model settings for enhanced performance. The performance of the model and data was evaluated by receiver operating characteristic (ROC) curve, area under the curve (AUC), accuracy, precision, recall and F1 score on different cutoffs of disease duration. The sensitivity, specificity, PPV, NPV, accuracy, and F1 score were calculated at the optimal cutoff value that maximized the Youden index. The model and data with the best performance were selected for further analysis.

2.5. feature selection

Too many features can lead to overfitting and reduce the interpretability of the model. Therefore, feature selection was performed on final model for further optimization. Although, feature importances can be evaluated directly from the boosted trees, these importances have been shown to be local and inconsistent. A SHAP score inspired by Shapley values can combine different explanation models and provide a global feature importance score that is consistent across different test sets. However, other than to arbitrarily select an importance threshold beyond which features are considered unimportant, SHAP analysis does not offer an algorithmic way to filter a large feature set to a limited set of important features. Boruta algorithm is a wrapper method that determine the importance of features by comparing the importance of original features with their shuffled copies. If importance of an original feature is significantly greater than its shuffled copy, that feature is deemed important (Kursa and Rudnicki, 2010).

In this study, Boruta algorithm was performed on the final model, with max iteration set to 50 and repetition for 15 times. The confirmed features list and importance rank were recorded for each run. The overlap of confirmed features across runs was calculated to determine the final feature list. For time dependent data, once a laboratory item was confirmed as important in any phase of disease, the feature of the same lab item in other phase of disease was also considered as important. That is to say, when top 10 items were selected, the features representing different phases of that item were all included in the final model, which is at most 30 features for time dependent data. The model performance on different number of top items, ranked by the average importance score, were compared to determine the optimal number of features for the final model.

2.6. Model explanation

SHAP (Shapley Additive Explanations) is an explanation method based on game theory, used to interpret the outputs of machine learning models. It assigns each feature's contribution to the model's prediction, helping users understand the decision-making process of the model. In tree models that are not sensitive to collinearity, such as Xgboost and GBM, the model tends to select one most important feature among collinear features for a split and ignores the other correlated features. As a result, SHAP may assign a high contribution to the first selected feature while the others receive little or no contribution. This happens because the model captures most of the information through the first feature, effectively overshadowing the others. By introducing joint distribution estimation for correlated features

during SHAP calculations, we can capture their dependencies and ensure that SHAP values reflect the true impact of correlated features (Lundberg and Lee, 2017). The shapley value of each laboratory item in different phase of disease were compared to reveal various predicting ability. Some of the features were further analyzed with different age populations to explore the potential age effect on the predicting ability of laboratory data.

3. Results

3.1. Patient characteristics

Of the 9921 patients who visited the urticaria specialty clinic and met the inclusion criteria from January 2018 to December 2024. 3954 patients were selected for time independent dataset and 1948 patients were selected for time dependent dataset, the rest were excluded due to serious missing data. 3954 patients in time independent data and 1948 patients time dependent data were allocated into separate derivation and external validation sets with a ratio of 7:3, respectively. The comparison of demographic and clinical variables among the training and external validation sets is shown in Table Appendix A for time independent dataset and Table Appendix A for time dependent dataset. No substantial differences were observed between the training set and the external test set across the majority of features in either time independent or time dependent data. Details of the study design are displayed in Figure 1.

Table 1 delineates the disparities between groups regarding the disease duration of patients in time independent Dataset. Among the 3954 patients, 2027 patients had a visit duration of less than 100 days, while 1927 patients had a visit duration of 100 days or more. The mean visit duration was 25.37 days for the short duration group and 604.32 days for the long duration group. The comparison of clinical characteristics between patients with short and prolonged visit duration in the time-independent dataset (Table 1) highlights several trends that align with findings from previous studies on chronic spontaneous urticaria (CSU). In this analysis, immunoglobulin E (IgE) (131.65 ± 242.00 in good outcome vs 152.28 ± 346.75 in poor outcome), eosinophils percentage (1.55 ± 2.02 in good outcome vs 2.31 ± 2.37 in poor outcome), eosinophils counts (128.60 ± 141.68 in good outcome vs 162.83 ± 190.02 in poor outcome), and basophil percentage (0.25 ± 0.22 in good outcome vs 0.32 ± 0.24 in poor outcome) were significantly higher in patients with poor outcomes, which is consistent with previous studies that have linked these markers to CSU severity and duration (SanchezBorges et al., 2017; Rabelo-Filardi et al., 2013; Kolkhir et al., 2019). In contrast, C-reactive protein (CRP) levels were significantly lower in patients with poor outcomes (13.03 ± 19.62 in good outcome vs 7.71 ± 10.23 in poor outcome), which contradicts previous findings that have associated higher CRP levels with CSU severity (Rabelo-Filardi et al., 2013). However, this discrepancy could reflect the complexity of CRP as an inflammatory marker that may vary across different disease stages or subgroups of patients. In time dependent data, which is shown in supplementary Table Appendix A, CRP presented various trends in different phases of disease, with significantly lower CRP levels in acute phase (10.21 ± 10.95 in good outcome vs 8.95 ± 8.22 in poor outcome) and higher CRP levels in chronic phase (5.39 ± 4.84 in good outcome vs 6.47 ± 8.81 in poor outcome) in patients with poor outcomes. This finding suggests that the predictive ability of CRP may be influenced by the disease phase, which could explain the inconsistent results observed in the time-independent dataset.

3.2. Comparison of multiple models on time-dependent and time-independent data

Time independent and Time dependent data, combined with 5 different models (XGBoost, Random Forest, AdaBoost, GBM, SVM), were used to generate 10 machine learning models to predict the disease duration of urticaria. Although most models utilized in this study is based on regression trees, their performance in directly predicting continuous numeric values in the regression task was suboptimal, which is likely due to the complexity of the underlying data distribution and model constraints. The predicted values, while not accurate for continuous outcome prediction, can still be utilized for binary classification tasks by applying an optimal threshold. Specifically, the model's output is better suited for classification of disease duration when applying a predefined threshold that maximizes the classification performance (e.g., based on ROC-AUC or Youden's index). Therefore we use binary classification task to evaluate the model performance. Among the 10 models, the XGBoost model with time-dependent data ($AUC = 0.906 \pm 0.013$ for 42, $AUC = 0.926 \pm 0.013$ for 100, $AUC = 0.884 \pm 0.018$ for 365) has the best predictive effect for disease duration, followed by the Random Forest model with time-dependent data ($AUC = 0.933 \pm 0.025$ for 42, $AUC = 0.922 \pm 0.028$ for 100, $AUC = 0.837 \pm 0.037$ for 365). The ROC curves for the top three best-performing ML models are presented in Figure 2. The discriminative performances and ROC curves of all 10 models are listed in Supplementary Table Appendix A, and Supplementary Figure Appendix A, respectively.

Characteristic	Good Outcome	Poor Outcome	P-value
Number of patients	2027	1927	
Outcome	25.37 ± 25.09	604.32 ± 437.04	0.000 ↑
Gender	0: 64.7%, 1: 35.3%	0: 61.1%, 1: 38.9%	0.023
First Visit Age	30.12 ± 21.79	28.11 ± 23.61	0.005 ↓
CI nd U	0: 99.3%, 1: 0.7%	0: 97.2%, 1: 2.8%	0.000
Lymphocytes Percentage	28.98 ± 14.05	34.02 ± 13.88	0.000 ↑
Neutrophils Percentage	63.13 ± 15.68	56.63 ± 14.97	0.000 ↓
Monocytes Percentage	6.18 ± 1.64	6.49 ± 1.67	0.000 ↑
Mean Corpuscular Hemoglobin Concentration	335.62 ± 10.51	336.93 ± 10.30	0.000 ↑
Platelet Count	263.47 ± 74.09	256.73 ± 68.96	0.003 ↓
White Blood Cell Count	10.31 ± 3.00	9.35 ± 2.39	0.000 ↓
Mean Corpuscular Hemoglobin	29.07 ± 2.33	28.98 ± 2.30	0.202
Mean Corpuscular Volume	86.66 ± 6.65	85.99 ± 6.67	0.002 ↓
Hemoglobin	130.03 ± 9.16	130.20 ± 8.46	0.546
Eosinophils Percentage	1.55 ± 2.02	2.31 ± 2.37	0.000 ↑
Basophils Percentage	0.25 ± 0.22	0.32 ± 0.24	0.000 ↑
Absolute Eosinophil Count	0.13 ± 0.20	0.18 ± 0.21	0.000 ↑
Absolute Lymphocyte Count	2.61 ± 1.12	2.83 ± 1.29	0.000 ↑
Mean Platelet Volume	9.14 ± 1.39	9.02 ± 1.39	0.007 ↓
Platelet Distribution Width	12.72 ± 1.96	12.57 ± 1.94	0.020
Eosinophil Count Absolute	128.60 ± 141.68	162.83 ± 190.02	0.000 ↑
CR eactive Protein	13.03 ± 19.62	7.71 ± 10.23	0.000 ↓
Immunoglobulin E	131.65 ± 242.00	152.28 ± 346.75	0.029
SMRNP	1.18 ± 2.16	1.22 ± 1.84	0.576
Anti SSA	1.42 ± 5.44	1.58 ± 5.68	0.363
Anti Jo 1	1.08 ± 1.77	1.12 ± 2.06	0.534
Nucleosome	0.57 ± 0.34	0.65 ± 0.43	0.000 ↑
Ribosomal PP rotein	1.07 ± 0.71	1.17 ± 2.13	0.050
Ro 52	2.10 ± 5.90	2.23 ± 6.81	0.515

Table 1: Comparison of the characteristics between patients with good and poor outcomes in the time independent dataset continuous variables are presented as mean ± standard deviation, categorical variables are presented as number (percentage) good outcome is defined as visit duration < 100 days, poor outcome is defined as visit duration ≥ 100 days

3.3. Feature selection and final model

Xgboost with time dependent data, as the best performing model, was selected for further optimization. Feature importance ranking by Boruta algorithm are shown in Figure 3 and Supplementary Table Appendix A. A total of 38 laboratory items were confirmed as important by Boruta algorithm, as shown in Supplementary Table Appendix A. During the process of reducing features based on the feature importance, the changes in AUCs for the model shows that the model maintain a robust performance until the number of items reduced to less than 10, as shown in Figure 4 A-C. The 25-item model shows the best performance in internal validation set with AUC of 0.914, 0.939, 0.899 for 42, 100, 365 cutoffs, respectively. The 25-item model were selected as the final model, and showed a stable performance in external test set with AUC of 0.887, 0.898, 0.809 for 42, 100, 365 cutoffs, as shown in Figure 3 4 D-F.

3.4. Model explanation

The SHAP analysis was performed on the final model with 25 top items. The heatmap shows shap values of the top 25 features on 300 samples from dataset, with the samples ordered by hierarchical clustering based on the similarity of shap values, and the corresponding model output and global importance of each feature, shown in Figure 5.

3.4.1. SHAP analysis reveals alignment with previous studies while discovering phase-specific trends

SHAP analysis on well-known markers of CSU severity and duration, such as eosinophils, basophils, and CRP, revealed consistent trends with previous studies while providing insights on how those markers in different phase can have different predicting ability. A high value in Basophils percentage during chronic phase of disease pushed the model output towards a longer visit duration, a trend that was proved by previous studies (SanchezBorges et al., 2017), while a high Eosinophils percentage during the preclinical phase of disease had the opposite effect, as shown in Figure 6 A. The increasing CRP levels in the chronic phase of disease was having positive effect on the model output, same as the previous studies that have associated higher CRP levels with CSU severity (Rabelo-Filardi et al., 2013), while the extreme high CRP levels in the acute phase of disease was having negative effect on the model output, as shown in Figure 6 B. High immunoglobulin E levels in the chronic phase of disease was having positive effect on the model output, which is consistent with previous studies that have linked high IgE levels to CSU severity and duration (SanchezBorges et al., 2017), while having no effect in the acute phase of disease, as shown in Figure 6 C. It has been discovered that lymphocytes count is reduced in acute urticaria compared with healthy control. But whether a lower lymphocytes count is associated with severity and duration of CSU is still unknown. SHAP analysis shows that a low lymphocytes count in the preclinical and acute phase is having positive effect on the model output, while having negative effect in the chronic phase, as shown in Figure 6 D.

Thyroid dysfunction is a common comorbidity in patients with chronic urticaria, and it has been reported that reducing thyroxine dose may cause flare-ups of chronic urticaria and angio-oedema, suggesting a possible link between thyroid function and urticaria severity (Dunkley and Jawad, 2003). Our study shows that a high TSH level in the chronic phase of disease is having negative effect on the model output, as shown in Figure 6 E.

3.4.2. SHAP analysis in different age groups shows novel age effect

By subsetting data into different age groups, we found that the predicting ability of laboratory data can be different in different age groups. For example, we observed that Absolute eosinophil count in the acute phase of disease is having positive effect on the model output in all data. When subsetting the data into different age groups, we found that the predicting ability of Absolute eosinophil count in the acute phase of disease is more significant in the age group of 0-2, and almost no effect when the patient is older than 6 years old, as shown in Figure 7. Another target showing age effect is Absolute lymphocytes count. In Shap analysis on all data, the a high lymphocytes count during acute and preclinical phase is beneficial while a high lymphocytes count during chronic phase is harmful. When subsetting the data into different age groups, we found that the predicting ability presented before is only preserved when patients are older than 6 years old, with no effect in the age group of 0-2, and little effect in the age group of 2-6, as shown in Figure 8.

4. Discussion

Although the pathogenesis of CSU is not yet fully understood, it is well established that its signs and symptoms are due to the activation of mast cells and basophils, leading to the release of histamine and other inflammatory mediators(Zuberbier et al., 2021). Based on recent evidence, it is known that the causes of CSU include autoimmunity Type I (CSUaiTI, or “autoallergic CSU”; with IgE autoantibodies to self-antigens) and autoimmunity Type IIb (CSUaiTIIb; with mast cell-directed activating autoantibodies). In CSU due to unknown cause (CSUuc), as of yet unknown mechanisms are relevant for the degranulation of skin mast cell (MC)(Sellal et al., 2023; Maronese et al., 2023). The results of the basic tests performed in CSU can point to CSUaiTI vs CSUaiTIIb, with CRP more often elevated and eosinophil and basophil levels more often reduced in CSUaiTIIb(Xiang et al., 2023).

Eosinopenia in chronic spontaneous urticaria patients is associated with type IIb autoimmunity, high disease activity, and poor treatment response. (Kolkhir et al., 2019), especially in children (A et al., 2023).

Acute urticaria patients have abnormal cell immune responses, with lower numbers of CD3+ and CD4+ lymphocytes compared to healthy controls.(De-yu, 2009).

Chronic idiopathic urticaria is associated with a prominent infiltrate of T-lymphocytes, monocytes, and mast cells, suggesting potential interactions between these cells to cause mediator release.(Elias et al., 1986)

Reducing thyroxine dose may cause flare-ups of chronic urticaria and angio-oedema, suggesting a possible link between thyroid function and musculoskeletal conditions.(Dunkley and Jawad, 2003)

Mean platelet volume levels are higher in patients with chronic urticaria and correlate with its severity, suggesting coagulation and inflammation may play a role in the disease.(Aleem et al., 2015) Chronic urticaria with a positive autologous serum skin test is associated with higher clinical severity, increased platelet volume, and increased C-reactive protein levels.(Magen et al., 2010)

Neutrophilic urticarial dermatosis is a distinct cutaneous manifestation of neutrophilic aseptic disease, strongly associated with systemic diseases like Schnitzler syndrome, adult-onset Still disease, lupus erythematosus, and hereditary autoinflammatory fever syndromes.(Kieffer et al., 2009) Neutrophilic urticaria is common and not associated with other diseases, but its presence in biopsy samples may indicate rheumatic disease.(Llamas-Velasco et al., 2012)

The SHAP analysis performed in this study provides valuable insights into the prediction of visit duration for chronic spontaneous urticaria (CSU) based on clinical laboratory data, aligning with previous studies while also revealing novel findings. Consistent with prior research, our model demonstrates that elevated basophil percentages during the chronic phase of CSU are associated with longer visit durations, in line with previous findings (SanchezBorges et al., 2017). Similarly, increased CRP levels during the chronic phase were found to have a positive effect on the model’s output, corroborating earlier studies that have linked higher CRP levels with CSU severity and prolonged disease duration (Rabelo-Filardi et al., 2013). However, an unexpected negative effect of extreme CRP levels during the acute phase was identified, indicating that elevated CRP might play a different role in the early stages of the disease. Immunoglobulin E, absolute lymphocytes count, free thyroxine, and mean platelet volume were also found to have significant impacts on the model output, consistent with previous research that has highlighted the importance of these markers in CSU severity and duration (SanchezBorges et al., 2017; Aleem et al., 2015; Dunkley and Jawad, 2003; Magen et al., 2010).

A novel discovery in our analysis is the phase-specific behavior of certain biomarkers. While a high eosinophil percentage in the preclinical phase was associated with shorter visit durations, the same marker in the chronic phase contributed to prolonged visits, highlighting distinct dynamics depending on the disease stage. Additionally, the SHAP analysis revealed that IgE levels, often associated with CSU severity and duration (SanchezBorges et al., 2017), only had a significant impact during the chronic phase, with no effect in the acute phase. This phase-specific difference in the predictive power of laboratory markers represents a significant innovation in understanding CSU’s progression.

Our study also uncovered age-specific trends, particularly regarding eosinophil and lymphocyte counts. While eosinophil levels were generally predictive of longer visit durations in the acute phase, this effect was markedly stronger in younger patients (0-2 years) and diminished in older age groups (above 6 years), aligning with previous research that has highlighted age-related variations in eosinophil levels (A et al., 2023).

It has been discovered that acute urticaria patients have lower numbers of CD3+ and CD4+ lymphocytes compared to healthy controls (De-yu, 2009). In our study, lymphocyte counts exhibited a dual-phase impact, being beneficial during the acute and preclinical phases but detrimental in the chronic phase. This trend was consistent only in older patients (above 6 years), with minimal effects observed in younger age groups. These age-related variations suggest

that laboratory markers may not uniformly predict CSU duration across all age groups, a finding that could inform personalized management strategies based on patient age.

In conclusion, our SHAP analysis both validates and extends previous knowledge of CSU biomarkers, particularly by identifying phase-specific and age-specific trends in their predictive power. These findings have the potential to enhance clinical decision-making by providing a more nuanced understanding of how different biomarkers contribute to CSU's disease course across various patient demographics.

5. code availability

Code for data extraction, model development, and model explanation are available at <https://github.com/jabberwockyang/UrticariaPrediction>

Appendix A. Supplementary data

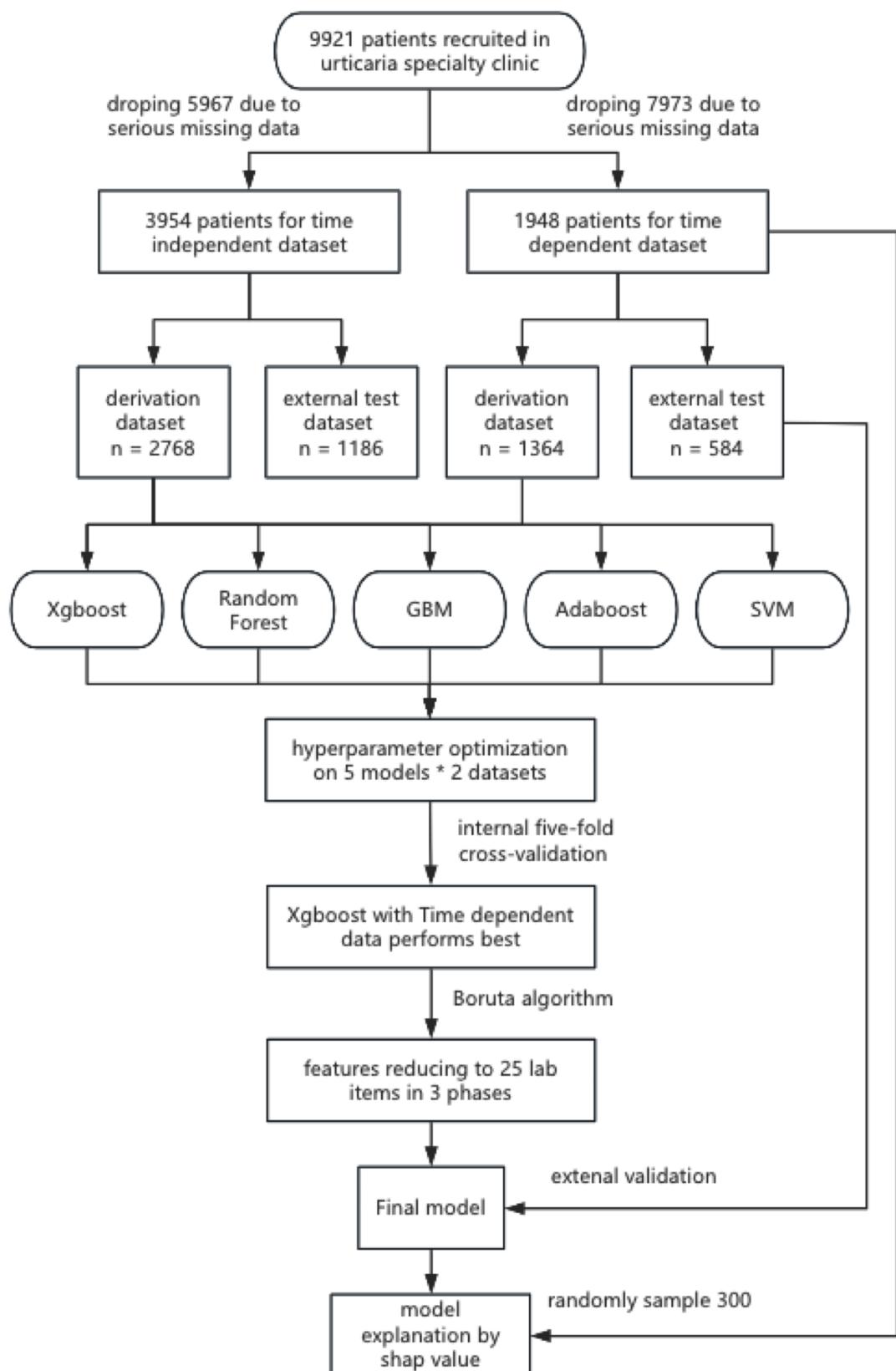
- Supplementary file 1: Database schema
- Supplementary file 2: SQL queries for time independent extraction
- Supplementary file 3: SQL queries for time dependent extraction
- Supplementary Table1: Comparison of the characteristic between the training set and external testing set data in the time independent dataset
- Supplementary Table2: Comparison of the characteristic between the training set and external testing set data in the time dependent dataset
- Supplementary Table 3: Comparison of the characteristic between the short duration group and long duration group in the time dependent dataset
- Supplementary Table 4: The discriminative performances of all 10 models in 5 fold cross validation
- Supplementary Figure 1: ROC curves of all 10 models in 5 fold cross validation
- Supplementary Table 5: Feature importance ranking by Boruta algorithm
- Supplementary Table 6: 38 laboratory items confirmed as important by Boruta algorithm
- Supplementary Table 7: The performance of the final model on external test set

Appendix B. Acknowledgements

References

- A, N.S., B, P, M, S., Chidambaranathan, S., Kaviarasan, P., 2023. Serum ige and absolute eosinophil count-a valuable serological markers in childhood urticaria. *IP Indian Journal of Clinical and Experimental Dermatology* doi:10.18231/j.ijced.2023.006.
- Aleem, S., Masood, Q., Hassan, I., 2015. Correlation of mean platelet volume levels with severity of chronic urticaria. *Journal of Dermatology and Dermatologic Surgery* 19, 9–14. doi:10.1016/J.JDDS.2014.03.006.
- Balp, M., Vietri, J., Tian, H., Isherwood, G., 2015. The impact of chronic urticaria from the patient's perspective: A survey in five european countries. *The Patient* 8, 551–558. doi:10.1007/s40271-015-0145-9.
- Chiang, H.L., Chen, C.H., Koo, M., Tsai, T.Y., Wu, C.H., 2022. Predictors of response to oral medications and low-histamine diet in patients with chronic urticaria. *Journal of Immunology Research* 2022. doi:10.1155/2022/5243825.
- Curto-Barredo, L., Archilla, L.R., Vives, G., Pujol, R., Giménez-Arnau, A., 2018. Clinical features of chronic spontaneous urticaria that predict disease prognosis and refractoriness to standard treatment. *Acta dermato-venereologica* 98 7, 641–647. doi:10.2340/00015555-2941.
- De-yu, C., 2009. Determination of peripheral blood t lymphocytes and their subsets in patients with acute urticaria. *China Medical Abstract of Dermatology* .
- Dunkley, L., Jawad, A.S.M., 2003. Thyroid function in chronic urticaria and angio-oedema. *Journal of the Royal Society of Medicine* 96 11, 571. doi:10.1258/JRSM.96.11.571.
- Elias, J., Boss, E., Kaplan, A., 1986. Studies of the cellular infiltrate of chronic idiopathic urticaria: prominence of t-lymphocytes, monocytes, and mast cells. *The Journal of allergy and clinical immunology* 78 5 Pt 1, 914–8. doi:10.1016/0091-6749(86)90240-X.
- Kieffer, C., Cribier, B., Lipsker, D., 2009. Neutrophilic urticarial dermatosis: A variant of neutrophilic urticaria strongly associated with systemic disease. report of 9 new cases and review of the literature. *Medicine* 88, 23–31. doi:10.1097/MD.0b013e3181943f5e.
- Kim, Y., Han, K.D., Bang, C., Lee, J.H., Lee, J.Y., Park, Y., Park, Y., 2021. High waist circumference rather than high body mass index may be a predictive risk factor for the longer disease duration of chronic spontaneous urticaria. *Scientific Reports* 11. doi:10.1038/s41598-021-81484-1.
- Kolkhir, P., Church, M., Altrichter, S., Skov, P., Hawro, T., Frischbutter, S., Metz, M., Maurer, M., 2019. Eosinopenia, in chronic spontaneous urticaria, is associated with high disease activity, autoimmunity and poor response to treatment. *The journal of allergy and clinical immunology. In practice* doi:10.1016/j.jaip.2019.08.025.

- Kursa, M.B., Rudnicki, W.R., 2010. Feature selection with the boruta package. *Journal of statistical software* 36, 1–13.
- Labbene, I., Tekou, A., 2023. Prevalence and treatment of chronic spontaneous urticaria in children. *Scholars Journal of Applied Medical Sciences* doi:10.36347/sjams.2023.v11i09.025.
- Lin, Y.R., Liu, T.H., Wu, T.K., Chang, Y.J., Chou, C., Wu, H.P., 2011. Predictive factors of the duration of a first-attack acute urticaria in children. *The American journal of emergency medicine* 29 8, 883–9. doi:10.1016/j.ajem.2010.04.004.
- Llamas-Velasco, M., Fraga, J., Requena, L., Sánchez-Pérez, J., Ovejero-Merino, E., García-Díez, A., 2012. [neutrophilic urticaria or urticaria with predominantly neutrophilic inflammatory infiltrate: study of its clinical and histopathologic characteristics and its possible association with rheumatic disease]. *Actas dermo-sifiliograficas* 103 6, 511–9. doi:10.1016/j.ad.2012.01.002.
- Lundberg, S., Lee, S., 2017. A unified approach to interpreting model predictions in 31st conference on neural information processing systems (nips 2017).(2017). CA, USA .
- Magen, E., Mishal, J., Zeldin, Y., Feldman, V., Kidon, M., Schlesinger, M., Sthoeger, Z., 2010. Increased mean platelet volume and c-reactive protein levels in patients with chronic urticaria with a positive autologous serum skin test. *The American Journal of the Medical Sciences* 339, 504–508. doi:10.1097/MAJ.0b013e3181db6ed5.
- Maronec, C., Ferrucci, S., Moltrasio, C., Lorini, M., Carbonelli, V., Asero, R., Marzano, A., Cugno, M., 2023. IgG and IgE autoantibodies to IgE receptors in chronic spontaneous urticaria and their role in the response to omalizumab. *Journal of Clinical Medicine* 12. doi:10.3390/jcm12010378.
- Poddighe, D., 2019. Letter to the editor: The prevalence of chronic spontaneous urticaria (csu) in the pediatric population. *Journal of the American Academy of Dermatology* doi:10.1016/j.jaad.2019.07.068.
- Pozderac, I., Lugović-Mihić, L., Artuković, M., Stipić-Marković, A., Kuna, M., Ferček, I., 2020. Chronic inducible urticaria: classification and prominent features of physical and non-physical types. *Acta dermatovenerologica Alpina, Pannonica, et Adriatica* 29 3, 141–148. doi:10.15570/actaaapa.2020.29.
- Rabelo-Filardi, R., Daltro-Oliveira, R., Campos, R., 2013. Parameters associated with chronic spontaneous urticaria duration and severity: A systematic review. *International Archives of Allergy and Immunology* 161, 197 – 204. doi:10.1159/000346896.
- Radonjić-Hoesli, S., Hofmeier, K., Micaletto, S., Schmid-Grendelmeier, P., Bircher, A., Simon, D., 2018. Urticaria and angioedema: an update on classification and pathogenesis. *Clinical Reviews in Allergy & Immunology* 54, 88–101. doi:10.1007/s12016-017-8628-1.
- Ring, J., Grosber, M., 2012. Urticaria: Attempts at classification. *Current Allergy and Asthma Reports* 12, 263–266. doi:10.1007/s11882-012-0275-2.
- SanchezBorges, M., Caballero-Fonseca, F., Capriles-Hulett, A., González-Aveledo, L., Maurer, M., 2017. Factors linked to disease severity and time to remission in patients with chronic spontaneous urticaria. *Journal of the European Academy of Dermatology and Venereology* 31. doi:10.1111/jdv.14221.
- Schettini, N., Corazza, M., Schenetti, C., Pacetti, L., Borghi, A., 2023. Urticaria: A narrative overview of differential diagnosis. *Biomedicines* 11. doi:10.3390/biomedicines11041096.
- Sella, J.A., Ferriani, M.P., Melo, J.M., Neto, O.T., Zanetti, M.E.T., Cordeiro, D.L., Lemos, Jos Eand Barros Jr, S.A., Aragon, D.C., Arruda, L.K., 2023. Type i and type iiB autoimmune chronic spontaneous urticaria: Using common clinical tools for endotyping patients with csu. *Journal of Allergy and Clinical Immunology: Global* 2, 100159.
- Xiang, Y., Guloglu, S., Elieh-Ali-Komi, D., Kocatürk, E., 2023. Chronic spontaneous urticaria: new evidences on the role of autoimmunity. *Current Opinion in Allergy and Clinical Immunology* 23, 438–445. doi:10.1097/ACI.0000000000000927.
- Zuberbier, T., Latiff, A.A.A., Abuzakouk, M., Aquilina, S., Asero, R., Baker, D., Ballmer-Weber, B., Bangert, C., Ben-Shoshan, M., Bernstein, J., Bindslev-Jensen, C., Brockow, K., Brzoza, Z., Neto, H.C.C., Church, M., Criado, P., Danilycheva, I., Dressler, C., Ensina, L., Fonacier, L., Gaskins, M., Gáspár, K., Gelincik, A., Giménez-Arnau, A., Godse, K., Gonçalo, M., Grattan, C., Grosber, M., Hamelmann, E., Hébert, J., Hide, M., Kaplan, A., Kapp, A., Kessel, A., Kocatürk, E., Kulthanian, K., Larenas-Linnemann, D., Lauferma, A., Leslie, T., Magerl, M., Makris, M., Meshkova, R., Metz, M., Micallef, D., Mortz, C., Nast, A., Oude-Elberink, H., Pawankar, R., Pigatto, P., Sisa, H.R., Gutiérrez, M.I.R., Saini, S., Schmid-Grendelmeier, P., Sekerel, B., Siebenhaar, F., Siiskonen, H., Soria, A., Staubach-Renz, P., Stingeni, L., Sussman, G., Szegedi, A., Thomsen, S., Vadász, Z., Vestergaard, C., Wedi, B., Zhao, Z., Maurer, M., 2021. The international eaaci/ga2len/euroguiderm/apaaaci guideline for the definition, classification, diagnosis, and management of urticaria. *Allergy* 77, 734–766. doi:10.1111/all.15090.



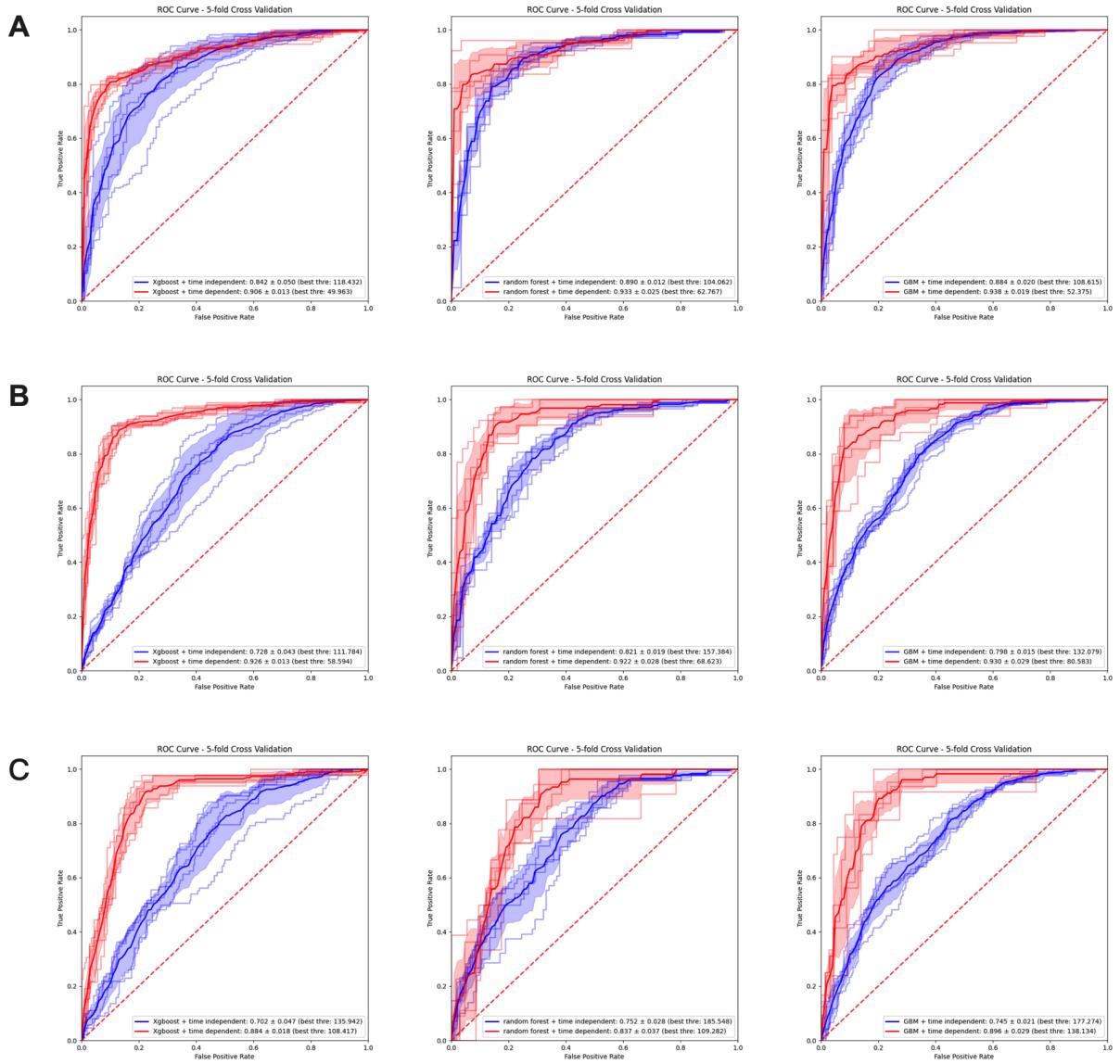


Figure 2: ROC curves for 5-fold cross-validation results for three different models (XGBoost, Random Forest, and GBM) at three binary classification thresholds (42, 100, and 365). Each row (A, B, and C) represents the results for a specific threshold. The models' performances are plotted with time-independent and time-dependent datasets, where red indicates the time-independent dataset and blue represents the time-dependent dataset. The shaded areas represent the 95% confidence intervals.

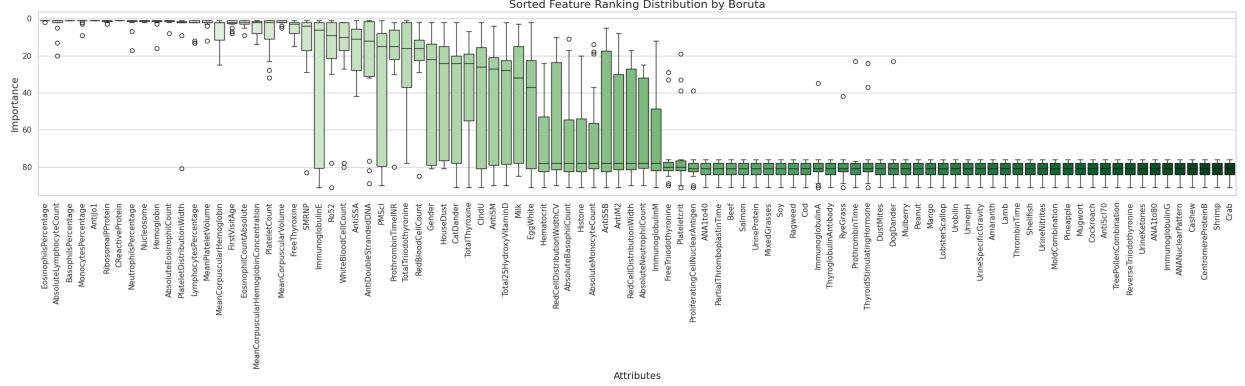


Figure 3: Boxplot representing averaged feature ranking distribution for each laboratory item generated by the Boruta algorithm over 15 iterations. The ranking value for each laboratory item is averaged for 3 different phases of disease. The y-axis shows the importance rank, with lower values indicating higher importance, while the x-axis lists the laboratory items names.

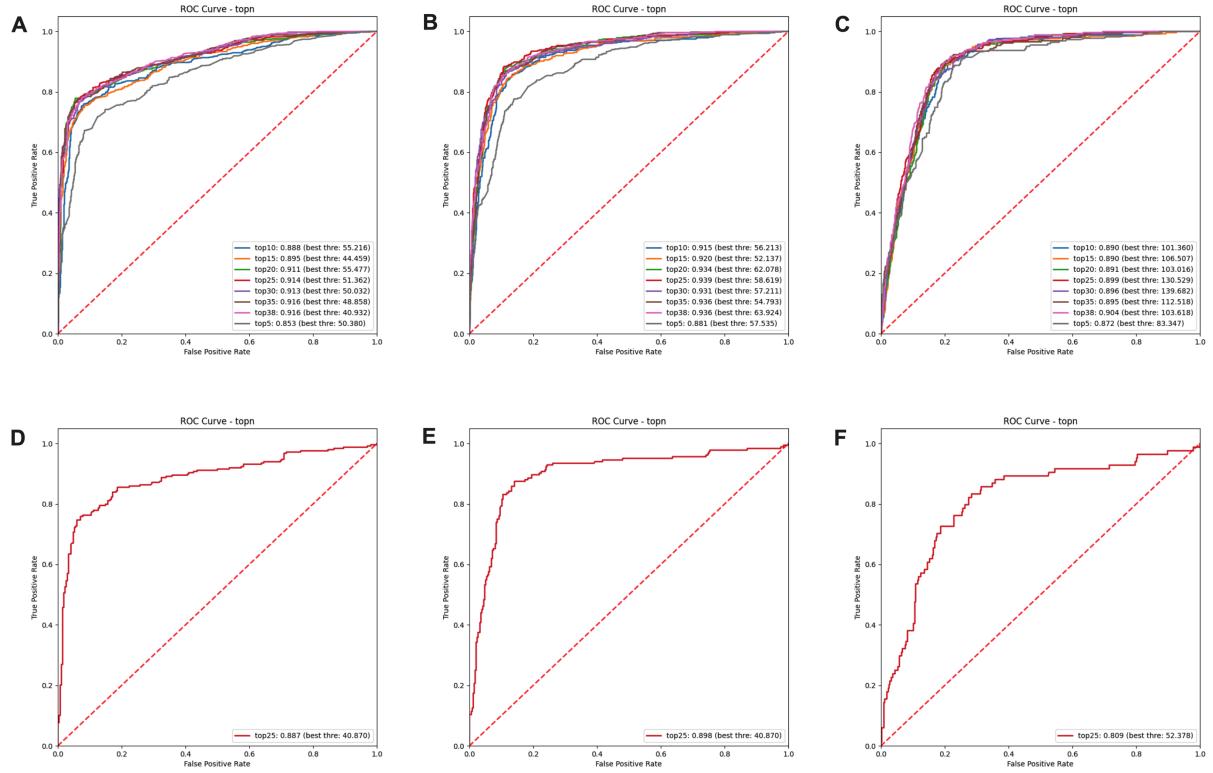


Figure 4: Panels A, B, and C display the ROC curves for the model as the number of top items is reduced based on the average importance score at cutoff values of 42, 100, and 365, respectively. Panels D, E, and F show the ROC curves for the final model with 25 top items on external test set at cutoff values of 42, 100, and 365, respectively.

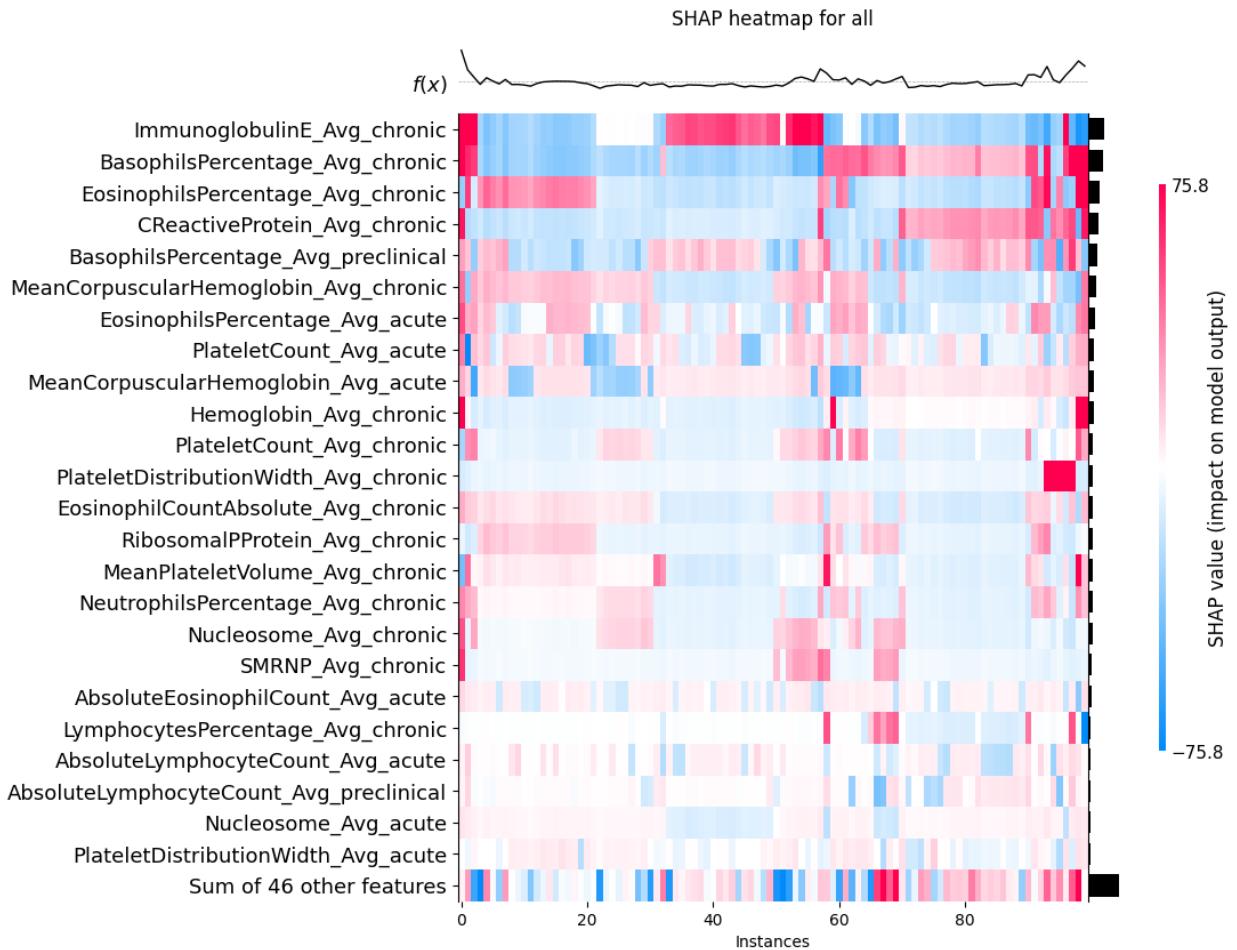
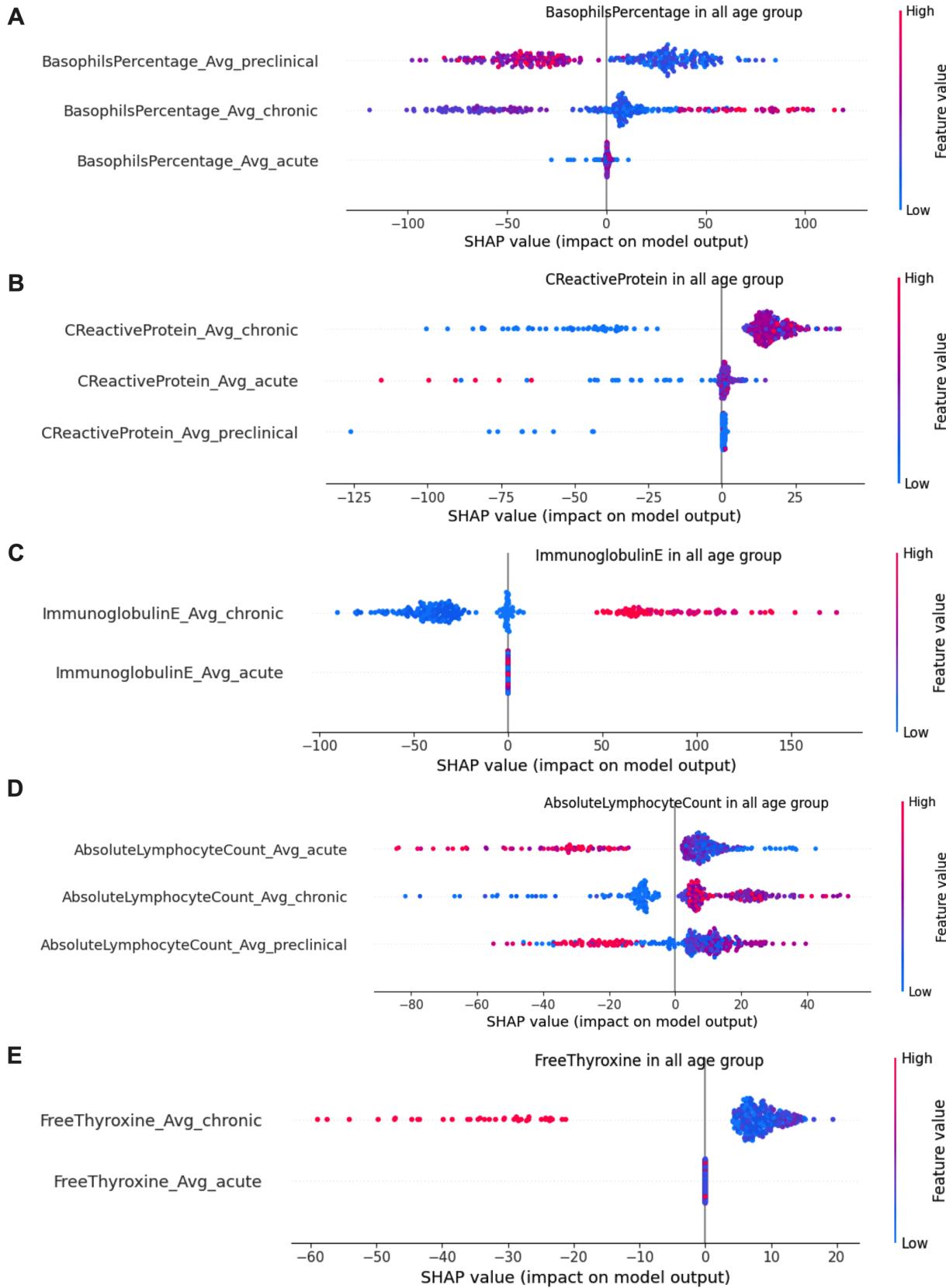


Figure 5: a heatmap plot with the instances on the x-axis, the model inputs on the y-axis, and the SHAP values encoded on a color scale. The samples are ordered based on a hierarchical clustering by their explanation similarity. This results in samples that have the same model output for the same reason getting grouped together. The output of the model is shown above the heatmap matrix, and the global importance of each model input shown as a bar plot on the right hand side of the plot.



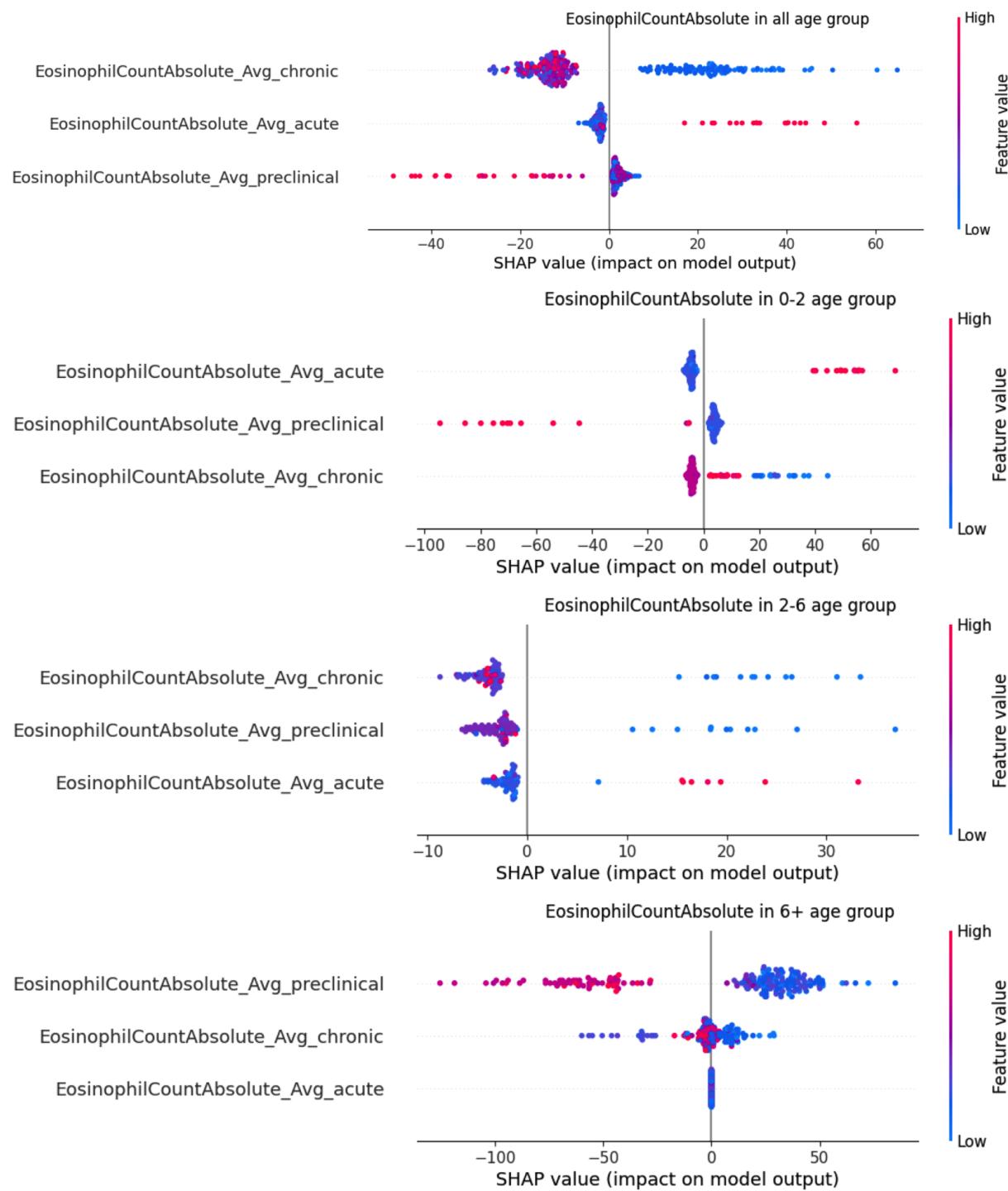


Figure 7:

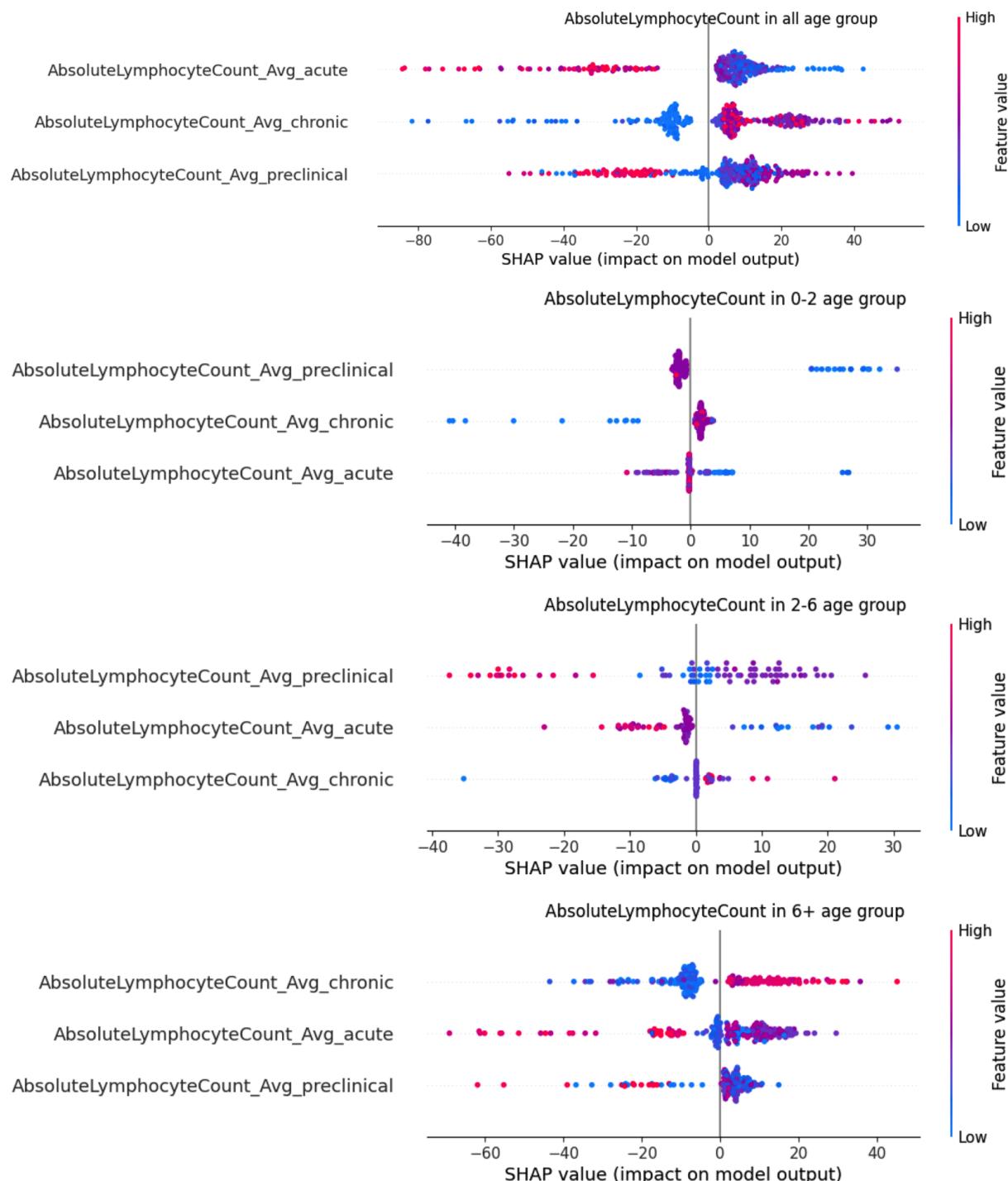


Figure 8: