Predictive Modeling of Behçet's Disease Using Machine Learning: Insights from Clinical Data Analysis

Abire ALLAOUI1235, Leila Barakat 4, Mina Moudatir 45, Khadija Echchilali 45, Hassan Elkabli 45
Internal medicine/rheumatology. Cheikh Khalifa University Hospital. Casablanca. Morocco1
Mohammed VI University of Health and Sciences, Casablanca, Morocco 2|Mohammed VI Center of Research and Innovation, Rabat, Morocco 3.|
Ibn Rochd University Hospital 4/Clinical Immunology, Autoimmunity and Inflammation Laboratory (LICIA), Faculty of Medicine and Pharmacy, Hassan II University, Casablanca, Morocco 5

Introduction:

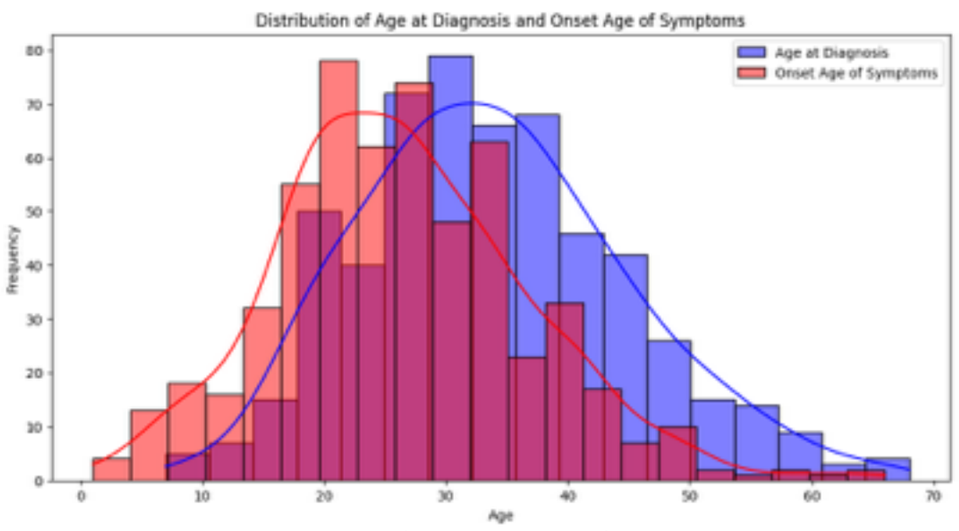
Behçet's disease (BD) is a chronic, multisystemic inflammatory disorder characterized by recurrent oral and genital ulcers, uveitis, and various systemic manifestations. The complexity of BD needs advanced methods like machine learning to better understand and predict disease patterns and outcomes.

Methods:

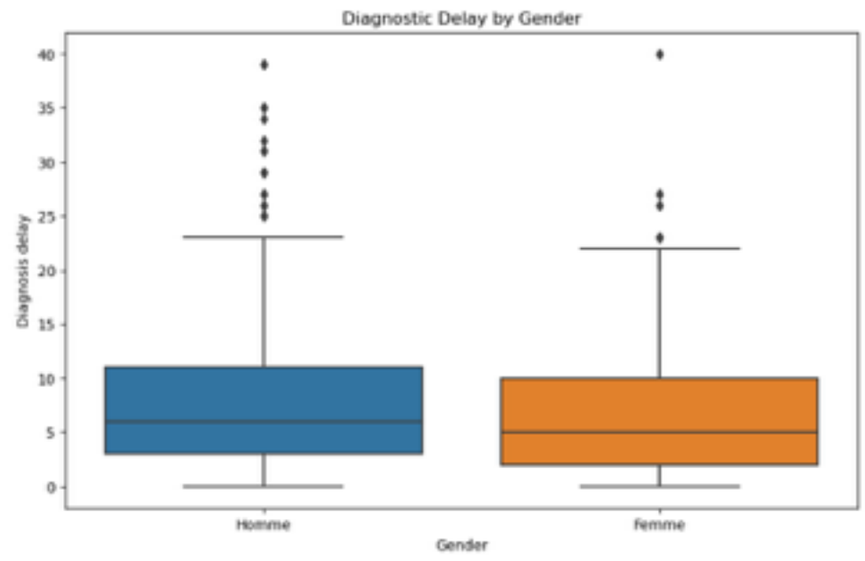
This study involves the analysis of clinical data from a cohort of BD patients using machine learning techniques. Data parameters include demographic information, clinical presentations, and disease progression. Statistical models were applied to identify key predictors of disease manifestations and outcomes.

Principal results:

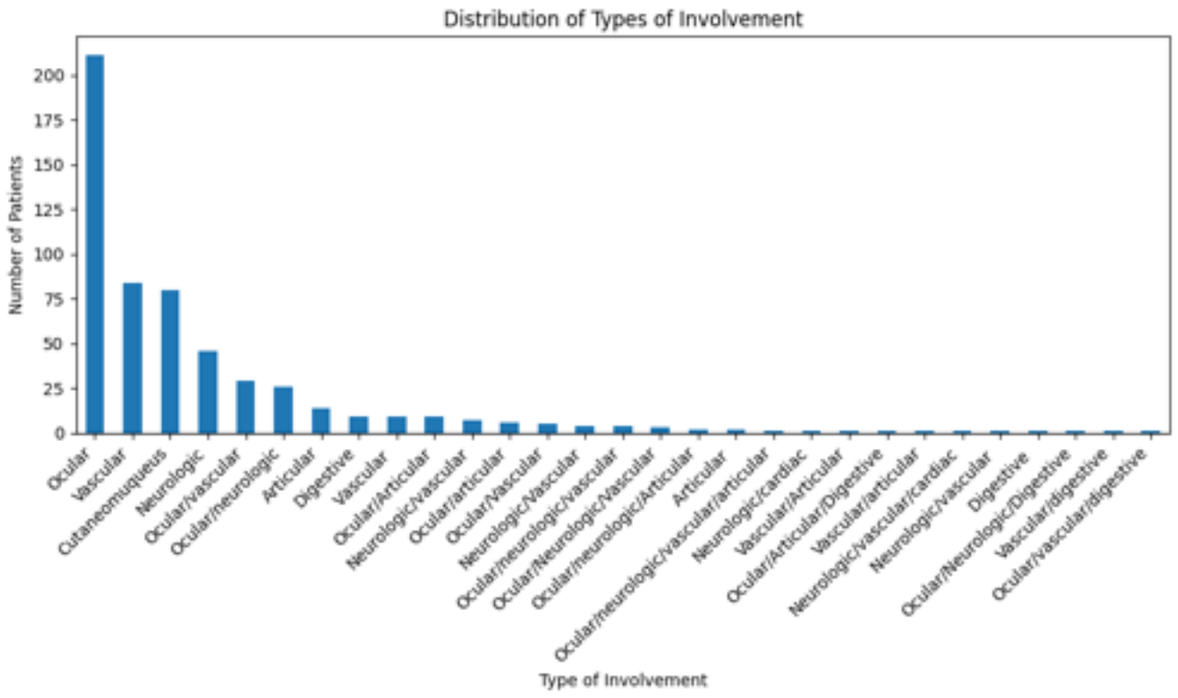
The cohort consisted of 561 patients (398 males and 163 females) with a mean age of diagnosis at 30 years. Machine learning analysis identified significant predictors of disease severity and organ involvement. For instance, the presence of familial history (OR = 2.3, 95% CI: 1.1-4.5) and cardiovascular risk factors (OR = 1.8, 95% CI: 1.0-3.2) were associated with increased systemic involvement. The analysis also highlighted the correlation between early onset of symptoms and more severe ocular involvement (AUC = 0.78).



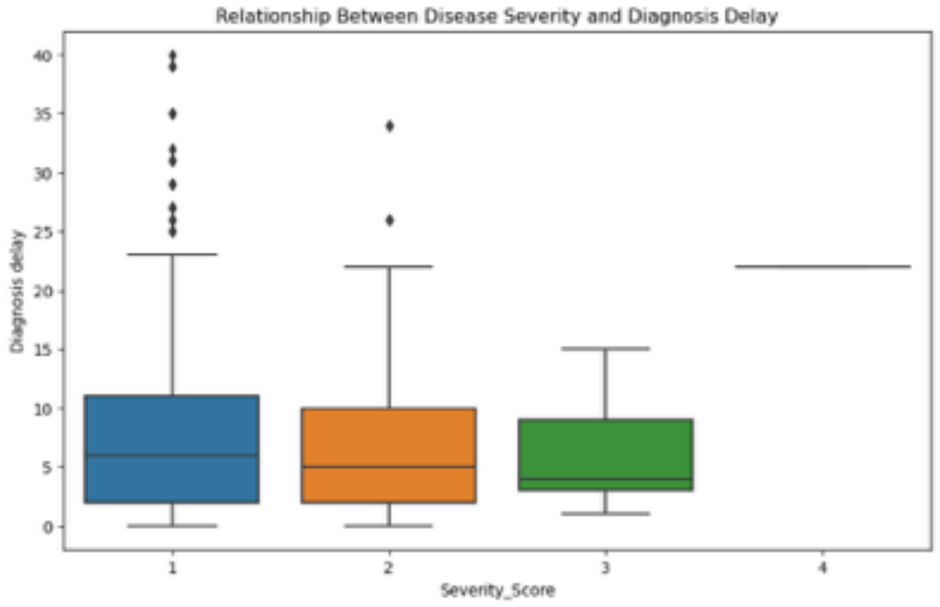
The graph presents a comparison of the distributions of two age-related variables: the age at diagnosis and the onset age of symptoms. The x-axis represents age, while the y-axis indicates the frequency of occurrences for each age range. The blue bars illustrate the distribution of the age at diagnosis, showing how patients' ages at the time of diagnosis are spread out. In contrast, the red bars represent the distribution of the onset age of symptoms, indicating at what age patients begin to exhibit symptoms. Superimposed on each histogram are kernel density estimation (KDE) curves, which provide a smoothed estimate of the probability density function for each variable, allowing for the observation of general trends and peaks in age.



By examining the spread and central tendency of the diagnostic delays for each gender, healthcare professionals can identify whether one gender experiences longer delays than the other. This information is crucial for understanding potential disparities in healthcare access and can inform strategies to improve diagnostic processes and reduce delays for all patients. Overall, this analysis contributes to a deeper understanding of the factors influencing diagnostic timelines in the patient population.



Visual representation of the distribution of different types of involvement among patients. Each bar in the chart indicates the frequency of a specific type of involvement, allowing for easy comparison across categories. The height of each bar reflects the number of patients associated with that type, highlighting which involvements are more prevalent. This information is crucial for understanding the condition's manifestations within the patient population and can inform clinical practices and treatment strategies



The analysis of disease severity provides insights into how the number of clinical manifestations relates to the delay in diagnosis. The box plot generated from the 'Severity_Score' and 'Diagnosis delay' variables visually represents this relationship. Each severity score reflects the total number of manifestations reported by a patient, with higher scores indicating more severe disease presentations. The y-axis shows the diagnosis delay, which is the time taken for patients to receive a diagnosis after the onset of symptoms. By examining the box plot, one can observe the distribution of diagnosis delays across different severity scores. If higher severity scores are associated with longer diagnosis delays, it may suggest that patients with more complex presentations face challenges in receiving timely diagnoses

Metric	Blindness	Death	Good evolution	Neurological Damage	Post-phlebitic Syndrome	Stabilization	Accuracy	Macro Avg	Weighted Avg
Precision	0.00	0.00	0.86	0.00	0.00	0.00	0.8496	0.14	0.75
Recall	0.00	0.00	0.98	0.00	0.00	0.00		0.16	0.85
F1-Score	0.00	0.00	0.92	0.00	0.00	0.00		0.15	0.80
Support	5	2	98	3	3	2	113	113	113

The results of the Random Forest classifier indicate an overall accuracy of approximately 85%. This means that the model correctly classified about 85% of the instances in the test set. However, a closer examination of the performance metrics reveals significant disparities across different classes.

Feature	Importance
Age of diagnosis	0.230927
Diagnosis delay	0.223563
Onset age of symptoms	0.211127
Oral Aphtosis flair per year	0.161690
Number of oral aphtosis per flair	0.136631
Gender_male	0.019352
Gender_female	0.016710

Interpretation of the Table:

- **Model Performance**: The model shows strong performance for the « Good evolution" class, with high precision (0.86), recall (0.98), and F1-score (0.92). However, it fails to predict the other classes, resulting in zero scores for precision, recall, and F1-score for "Blindness," "Death," "Neurological Damage," "Post-phlebitic Syndrome," and "Stabilization."
- **Overall Accuracy**: The overall accuracy of the model is approximately 85%, indicating that it correctly classifies 85% of the instances in the test set.
- **Macro and Weighted Averages**: The macro average scores are low, reflecting the model's poor performance on less frequent classes. The weighted averages are higher, indicating better performance when accounting for the number of instances in each class.
- **Feature Importance**: The feature importance analysis shows that "Age of diagnosis," "Diagnosis delay," and "Onset age of symptoms" are the most influential features in predicting the clinical course, while gender features have minimal impact.

This table format provides a clear and concise overview of the model's evaluation metrics and feature importance, facilitating easier interpretation and analysis.

Conclusions:

Machine learning models can effectively predict disease patterns in Behçet's disease, aiding in early diagnosis and personalized treatment plans.

Further research with larger cohorts is needed to validate these findings and enhance predictive accuracy.