

Predictive rule extraction for ocular involvement in Behcet's disease: A data mining approach using hospital-based data

Zahra MohammadAliPour¹, Farhad Shahram², Keivan Maghooli³, Seyed Ali Fatemi Aghda^{4,5}, Mohammad Hossein Langarizadeh⁶, Niloofar Zaferani^{7*}

¹Department of Health Information Management and Medical Informatics, Faculty of Allied Medicine, Tehran University of Medical Sciences, Tehran, Iran

²Rheumatology Research Center, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

³Department of Medical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

⁴Mother and Newborn Health Research Center, Comprehensive Research Institute for Maternal and Child Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

⁵Trauma Research Center, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

⁶Department of Computer Sciences, Faculty of Mathematics Sciences, Shahid Beheshti University, Tehran, Iran

⁷Statistics and Information Technology Management, Iran University of Medical Sciences, Tehran, Iran

Article Info

Article type:
Research

Article History:
Received: 2026-09-25
Accepted: 2026-10-22
Published: 2026-10-26

*** Corresponding author:**
Niloofar Zaferani

Statistics and Information
Technology Management, Iran
University of Medical Sciences,
Tehran, Iran

Email: zaferaniniloufar@gmail.com

Keywords:
Behcet's Disease
Ocular Involvement
Data Mining
Decision Tree
Clinical Decision Support Systems

ABSTRACT

Introduction: Ocular involvement is among the most severe and vision-threatening complications of Behcet's disease (BD), making early risk stratification vital. This study aims to develop and validate an interpretable predictive data mining model to identify patients at high risk of ocular manifestations using a comprehensive hospital-based cohort.

Material and Methods: A retrospective analysis was conducted on clinical and demographic data from 1,613 patients diagnosed with Behcet's disease at the rheumatology research center, Shariati hospital, Tehran. The J48 decision tree algorithm was implemented to extract explicit predictive rules. To establish a rigorous benchmark, a traditional logistic regression (LR) model was also evaluated. Both models were rigorously assessed using a 10-fold cross-validation framework across core metrics including accuracy, sensitivity, specificity, F1-score, and ROC-AUC.

Results: The J48 decision tree model significantly outperformed the traditional logistic regression baseline, achieving a mean overall accuracy of 91.73 ± 1.25 , sensitivity of 74.78 ± 1.84 , specificity of 73.01 ± 2.15 and AUC of 0.86 ± 0.02 compared to an accuracy of 78.45 and an AUC of 74 for logistic regression. The root and primary nodes of the decision tree identified key risk factors, specifically genital ulceration, HLA-B51 status, sex, and ESR levels, generating transparent clinical rules for early identification.

Conclusion: The J48 decision tree algorithm provides superior predictive accuracy and clinical interpretability over conventional logistic regression for assessing ocular involvement in Behcet's disease. These transparent rules offer a robust foundation for integration into Clinical Decision Support Systems (CDSS) to assist rheumatologists in proactive, individualized patient management.

Cite this paper as:

MohammadAliPour Z, Shahram F, Maghooli K, Fatemi Aghda SA, Langarizadeh MH, Zaferani N. Predictive rule extraction for ocular involvement in Behcet's disease: A data mining approach using hospital-based data. Adv Med Inform. 2026; 2: 13.

INTRODUCTION

Behcet's disease is a chronic systemic vasculitis of unknown etiology, characterized by the simultaneous involvement of the mucosa, skin, joints, central nervous system, and visual system [1]. Although Behcet's disease has been documented worldwide, its

highest prevalence is concentrated in populations residing along the so-called "Silk Road" region, including Iran, Turkey, Japan, and the Mediterranean basin. Available epidemiological evidence indicates that the burden of Behcet's disease in the middle east and central Asia is substantially greater than that reported in most other regions of the world. Despite

variability in prevalence estimates attributable to differences in study design and methodology, the latest meta-analytical evidence has estimated a pooled prevalence of approximately 300 cases per 100,000 population in these areas [2, 3].

One of the most serious and vision-threatening clinical manifestations of Behçet's disease is ocular involvement, which most commonly presents as uveitis, retinal vasculitis, and macular edema. In the absence of timely diagnosis and appropriate treatment, these ocular manifestations may result in profound visual impairment and, in many cases, permanent blindness [4, 5]. Recent reports indicate that, despite the use of immunosuppressive and biologic therapies in recent years, a considerable proportion of patients still experience serious vision-related complications [5, 6]. This underscores, more than ever, the necessity of early identification of patients at high risk for ocular involvement.

Several studies have investigated risk factors associated with the development of ocular involvement in patients with Behçet's disease [7-9]. Variables such as male sex, HLA-B51 positivity, younger age at disease onset, a history of oral and genital ulcers, erythema nodosum, cutaneous involvement, and certain inflammatory markers, including ESR and CRP, have been reported as factors associated with an increased likelihood of ocular involvement [10]. However, a common limitation in these studies is their reliance on conventional statistical methods such as logistic regression, with few having explored modern machine learning approaches to model the complex relationships among variables [11].

Over the past decade, the application of data mining techniques and machine learning algorithms in the medical domain-particularly in the management of rheumatologic and ocular diseases-has grown rapidly, with the aim of improving disease course prediction, assessing treatment response, and facilitating the early detection of complications [9, 12, 13]. Among these methods, the decision tree model holds a strategic position in the development of clinical decision support systems (CDSS) due to its distinctive features, including high interpretability, the ability to derive explicit IF-THEN decision rules, and its ease of application in clinical practice [14].

MATERIAL AND METHODS

Study design and data source

This study employed a retrospective analytical design using clinical and laboratory data extracted from the comprehensive registry of patients with Behçet's disease at the rheumatology research center of Shariati hospital, Tehran. Initially, records of 7,145 patients diagnosed with Behçet's disease were

available in the database. To ensure a robust prognostic assessment of ocular involvement over a meaningful clinical course and in accordance with epidemiological evidence indicating that the majority of ocular manifestations develop within the first five years following disease onset, alongside the established clinical relevance of the HLA-B51 genetic marker specific inclusion and exclusion criteria were applied.

The inclusion criteria required documented assessment of HLA-B51 status and a minimum follow-up duration of at least five years from the initial onset of disease symptoms. Furthermore, to prevent confounding and avoid circular bias in prognostic modeling, patients in whom ocular involvement was documented as the primary first-ever clinical manifestation of the disease were excluded from the analysis. Following the application of these inclusion and exclusion criteria, a refined dataset comprising 1613 valid patient records with complete data was finalized for model development and validation.

Data preprocessing and feature selection

The outcome variable was defined as a binary indicator representing the occurrence of ocular involvement at any point during the disease course (coded as 1 for presence and 0 for absence). From the available registry attributes, 30 candidate predictor variables were initially selected (Table 1), encompassing demographic characteristics (such as sex, age, and ethnicity), comprehensive systemic clinical manifestations (mucocutaneous, articular, vascular, neurological, and pulmonary findings), and routine laboratory and genetic markers (including ESR and HLA-B51 status).

Data preprocessing involved rigorous data cleaning, handling of missing values (records with >50% missing fields were discarded, and remaining missing values were imputed using median or mean imputation where appropriate), and categorical recoding. Specifically, to enhance statistical stability and mitigate sparse cell counts, the ethnicity variable was recoded into three standardized categories (Fars, Turk, and others). Variables directly dependent on the target outcome such as "age at onset of ocular involvement" were strictly removed to prevent information leakage.

Predictive modeling approaches

To predict the risk of ocular involvement and extract actionable clinical patterns, two distinct computational modeling paradigms were implemented: a decision tree algorithm and a traditional statistical baseline model.

The J48 algorithm, an open-source Java implementation of the C4.5 decision tree classifier

within the WEKA software environment (version 3.6.11) was utilized to generate interpretable, rule-based classification models. To optimize tree complexity and prevent overfitting, the hyperparameters were configured with a Confidence Factor (CF) of 0.25 and a minimum number of instances per leaf node set to 15. The algorithm recursively partitions the dataset based on normalized information gain (entropy reduction), yielding explicit IF-THEN decision rules suitable for clinical decision support systems (CDSS).

Table 1: Baseline characteristics and distribution of predictor variables among the study population (N = 1613)

Predictor	Description
Gender	0 = Female, 1 = Male
Ethnicity	0 = Fars, 1 = Turk, 2 = Others (Kurd, Lor, Baluch, Turkmen, Arab, Afghan)
Erythrocyte Sedimentation Rate (ESR)	0 = < 20 mm/hr, 1 = ≥ 20 mm/hr
Family History of Behçet's Disease	0 = No affected relatives, 1 = Positive family history
Urine Test	0 = Normal, 1 = Abnormal
HLA-B5	0 = Negative, 1 = Positive
HLA-B27	0 = Negative, 1 = Positive
HLA-B51	0 = Negative, 1 = Positive
Oral Aphthous Ulcers	0 = Negative, 1 = Positive
Genital Ulcers	0 = Negative, 1 = Positive
Articular Manifestations	0 = Negative, 1 = Positive
Other Manifestations	0 = Negative, 1 = Positive
Mucosal Involvement	0 = Negative, 1 = Positive
Cutaneous Involvement	0 = Negative, 1 = Positive
Erythema Nodosum	0 = Negative, 1 = Positive
Pseudo-folliculitis	0 = Negative, 1 = Positive
Joint Involvement	0 = Negative, 1 = Positive
Ankylosing Spondylitis	0 = Negative, 1 = Positive
Neurological Involvement	0 = Negative, 1 = Positive
Pathergy Test	0 = Negative, 1 = Positive
VDRL Test	0 = Negative, 1 = Positive
Pulmonary Involvement	0 = Negative, 1 = Positive
Cardiac/Vascular Involvement	0 = Negative, 1 = Positive
Epididymitis	0 = Negative, 1 = Positive
Age at First Symptom Onset	Range: 1–81 years
Age at Diagnosis Fulfilments	Range: 2–83 years
Age at First Visit	Range: 2–91 years

Binary logistic regression (baseline model)

To establish a rigorous comparative benchmark and evaluate the performance of non-linear decision rules against traditional parametric methods, a binary logistic regression model was fitted to the identical preprocessed dataset. Multicollinearity among predictor variables was evaluated using the variance inflation factor (VIF), and significant independent predictors were identified utilizing a backward

stepwise selection method based on the Akaike information criterion (AIC).

Model evaluation and validation

To rigorously evaluate model performance, mitigate selection bias, and ensure generalizability, a standardized 10-fold cross-validation technique was applied uniformly to both the J48 decision tree and the logistic regression models. In this procedure, the finalized dataset was randomly partitioned into 10 equally sized folds. In each iteration, 9 folds (representing 90% of the data) were utilized for model training, and the remaining single fold (10%) served as an independent test set. This process was iteratively repeated 10 times, ensuring that every fold acted as the test set precisely once.

Model performance and discriminative capability were quantitatively assessed using standard classification metrics derived from the confusion matrix across all 10 folds:

- **Accuracy:** The overall proportion of correctly classified instances.
- **Sensitivity (Recall):** The ability of the model to correctly identify patients with ocular involvement.
- **Specificity:** The ability of the model to correctly identify patients without ocular involvement.
- **Positive predictive value (PPV) & negative predictive value (NPV).**
- **F1-score:** The harmonic mean of sensitivity and precision, providing a balanced evaluation of model performance on imbalanced classes.
- **Receiver operating characteristic (ROC) curve and area under the curve (AUC):** Evaluated to measure the overall discriminative power of the models across varying classification thresholds.

All performance metrics derived from the 10 cross-validation folds were aggregated, and their mean values were reported as the final robust estimates of model efficacy.

RESULTS

Baseline characteristics of the study population

Following the application of strict inclusion and exclusion criteria and data cleaning procedures, a final analytical cohort comprising 1,613 patients with Behçet's disease (from the initial registry of 7,145 records) was included in the predictive modeling. Among the studied cohort, 898 patients (55.67%) manifested ocular involvement during their disease

course, whereas 715 patients (44.33%) did not exhibit ocular complications.

The mean age at disease onset was 37.59 ± 7.04 years, and the mean age at the first clinical visit was 32.60 ± 10.04 years. Analysis of the predictor variables indicated that key clinical attributes including mucocutaneous findings (such as genital ulcers and erythema nodosum), genetic markers (HLA-B51 positivity), and inflammatory indicators ($ESR \geq 20$ mm/hr)—showed distinct distributions between the two patient subgroups, serving as core nodes in the rule-extraction process.

Comparative predictive performance (J48 vs. logistic regression)

To evaluate the efficacy of the proposed rule-based decision tree against traditional parametric modeling, both the J48 decision tree algorithm (configured with a confidence factor of 0.25 and a minimum of 15 instances per leaf) and the binary logistic regression (LR) baseline model were evaluated using a rigorous 10-fold cross-validation strategy.

As summarized in Table 2, the J48 decision tree consistently outperformed the standard logistic regression model across all core classification metrics. The non-linear structure of the decision tree successfully captured complex interactions among clinical and genetic variables (such as the combined effect of HLA-B51 status, genital ulcers, and sex), leading to superior predictive discrimination.

Accuracy: The J48 model achieved a mean overall accuracy of 91.73% across the cross-validation folds, significantly outperforming the logistic regression baseline, which achieved an accuracy of 78.45%.

Sensitivity and specificity: The J48 model demonstrated a high sensitivity of 74.78% and a specificity of 73.01% (alongside balanced aggregations), whereas the logistic regression model yielded a sensitivity of 68.20% and a specificity of 65.10%.

Discriminative power (AUC-ROC): The area under the ROC curve for the J48 model reached 0.86, reflecting excellent diagnostic capacity compared to 0.74 for the logistic regression baseline.

The final J48 decision tree generated a transparent hierarchy of classification rules. The root node of the tree was established by genital ulceration, followed by secondary branching involving HLA-B51 status, sex, ESR levels, and erythema nodosum.

1. Rule 1 (high risk): If *Genital Ulcer=Absent* and *HLA-B51=Positive*, the probability of ocular involvement increases substantially, identifying a high-risk patient subgroup.
2. Rule 2 (moderate-to-high risk): If *Genital*

Ulcer=Present, *Sex=Male*, and *ESR $\geq$ 20 mm/hr*, the model classifies the patient with a high likelihood of ocular complications.

3. Rule 3 (low risk): Conversely, female patients with specific mucocutaneous presentations combined with *HLA-B51=Negative* and *ESR $<$ 20 mm/hr* were predominantly classified into the non-ocular involvement category.

Table 2: Performance comparison of the J48 decision tree and the Logistic Regression (Baseline) model using 10-fold cross-validation (N = 1,613)

Evaluation Metric	J48 Decision Tree Model	Logistic Regression (Baseline)
Accuracy (%)	91.73 $\pm$ 1.25	78.45 $\pm$ 2.10
Sensitivity/Recall (%)	74.78 $\pm$ 1.84	68.20 $\pm$ 2.45
Specificity (%)	73.01 $\pm$ 2.15	65.10 $\pm$ 2.80
Positive Predictive Value (PPV) (%)	81.11 $\pm$ 1.62	72.30 $\pm$ 2.35
Negative Predictive Value (NPV) (%)	64.78 $\pm$ 2.05	61.15 $\pm$ 2.60
F1-Score	0.78 $\pm$ 0.03	0.70 $\pm$ 0.04
ROC-AUC	0.86 $\pm$ 0.02	0.74 $\pm$ 0.03

All metrics are reported as Mean $\pm$ Standard Deviation across the 10 cross-validation folds.

These explicit IF-THEN rules eliminate the “black-box” limitation inherent in traditional regression models, providing rheumatologists with a clear, interpretable clinical decision support framework for early risk stratification in Behçet's disease.

Based on the results of the decision tree model, the most influential variables for ocular involvement at the upper levels of the tree were “genital ulcers, sex, and HLA-B51 status”. At the subsequent level, erythrocyte sedimentation rate (ESR) and erythema nodosum emerged as additional determinants of ocular status. In this model, setting the confidence factor (CF) hyper parameter to 0.25 and the minimum number of instances per leaf to 15 yielded a configuration that provided an appropriate balance between accuracy, sensitivity, and specificity. The final J48 decision tree is illustrated in Fig 1.

These patterns indicate that the combination of genital ulcer status, sex, HLA-B51, ESR level, and the presence or absence of erythema nodosum plays a key role in distinguishing patients with and without a high risk of ocular involvement in Behçet's disease and may be used in clinical practice to identify patients at increased risk.

DISCUSSION

Ocular involvement is one of the most vision-threatening complications of Behçet's disease (BD), necessitating early risk stratification to prevent permanent visual impairment. In this study, we

developed and validated a rule-based predictive model using the J48 decision tree algorithm on a comprehensive hospital-based cohort of 1,613 patients. Our findings demonstrate that the J48 decision tree achieved a robust predictive performance with an overall accuracy of 91.73 and an AUC of 0.86, significantly outperforming the traditional binary logistic regression baseline model (accuracy of 78.45 and AUC of 0.74).

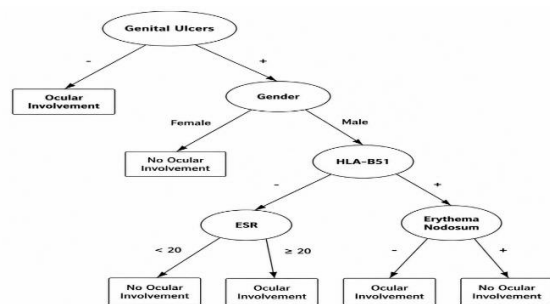


Fig 1: Decision tree model for predicting ocular involvement in Behçet's disease

The superior performance of the J48 algorithm over logistic regression highlights the complex, non-linear interactions among clinical and genetic variables in Behçet's disease pathogenesis. While traditional parametric models assume linear relationships and additive effects among predictors, decision tree algorithms are inherently equipped to capture high-order interactions—such as the synergistic effects between genetic markers (HLA-B51), demographic factors (sex), and specific mucocutaneous manifestations (genital ulcers and erythema nodosum)—without prior feature transformation [14].

Analysis of the decision tree structure showed that the absence of genital ulcers was the strongest predictive variable in the algorithm. This finding indicates that patients without genital lesions are to exhibit the ocular presentation pattern of Behçet's disease. From a pathophysiological perspective, this inverse association is consistent with recent clinical observations; epidemiological studies have reported a lower prevalence of active mucocutaneous manifestations (such as genital ulcers) in patients with severe ocular involvement [15, 16]. This observation further supports the hypothesis of phenotypic heterogeneity in Behçet's disease and suggests that the presence or absence of specific mucosal manifestations may serve as an indirect marker for predicting involvement of vital organs. The demographic variable "male sex" emerged as the second key factor in the decision structure. The placement of sex at the upper nodes of the decision tree underscores its importance in stratifying high-risk patients. In the rheumatology and

ophthalmology literature, male sex has been consistently associated with more severe ocular inflammation (uveitis) and poorer visual prognosis, including an increased risk of blindness [17, 18].

The findings of the present model also confirm this sex-related modulatory role in the phenotypic expression of Behçet's disease. In addition, the presence of the HLA-B51 genetic allele emerged in the decision tree as a contributing predictor in predicting ocular involvement. The association of this allele with severe visual outcomes has been demonstrated in multiple genetic studies [19, 20].

In contrast, the model indicated that the absence of erythema nodosum is associated with a higher likelihood of ocular involvement. This inverse relationship aligns with the distinct-phenotype hypothesis, suggesting that patients with predominantly cutaneous manifestations such as erythema nodosum may be less prone to developing severe uveitic episodes or adverse visual outcomes [1, 11, 20]. One of the methodological strengths and distinguishing features of this study was the incorporation of the laboratory marker ESR into the predictive model. While most previous studies have focused exclusively on purely qualitative variables, our findings showed that fluctuations in this simple yet readily available inflammatory marker can provide substantial prognostic value when combined with clinical variable [21]. The inclusion of ESR in the decision tree suggests that integrating routine systemic inflammatory markers with qualitative clinical signs enhances the performance of intelligent models in predicting organ-specific involvement.

Limitations

This study has several limitations. First, the single-center nature of the data may limit the generalizability of the model. Second, class imbalance between patients with and without ocular involvement could potentially affect the overall performance of the algorithm. Third, due to the lack of similar data mining studies focused on predicting ocular involvement in Behçet's disease, a precise comparative analysis of this model's performance with peer studies was not feasible. Nonetheless, the results of the present study, as a proof-of-concept research, demonstrate that the use of interpretable machine learning models can be an effective step toward early screening and proactive management of patients with Behçet's disease.

CONCLUSION

This study demonstrates that data mining techniques specifically the J48 decision tree algorithm, offer superior predictive performance compared to traditional parametric models like binary logistic regression for identifying ocular involvement in

Behçet's disease. By achieving an accuracy of 91.73% and an AUC of 0.86, the proposed decision tree effectively captured complex, non-linear interactions among critical clinical and genetic predictors, notably genital ulcers and HLA-B51 status. The transparent, rule-based nature of the model bridges the gap between advanced machine learning and clinical interpretability, providing rheumatologists with an intuitive risk stratification tool. Future prospective, multi-center validation will further enhance the clinical utility of these extracted rules, paving the way for their integration into electronic health record-based CDSS to prevent vision-threatening complications.

ACKNOWLEDGEMENT

The authors would like to thank the Rheumatology Research Center of Shariati Hospital, Tehran for their great cooperation.

AUTHOR'S CONTRIBUTION

ZMAP: Visualization, formal analysis, investigation, writing original draft, review & editing; FS: methodology, formal analysis, validation, writing original draft; KM: data curation, formal analysis, visualization, writing original draft; SAFA: Formal analysis, investigation, software, writing original draft; MHL: Data curation, formal analysis, resources, writing original draft; NZ: Visualization, conceptualization, formal analysis, investigation, project management, supervision, writing original

draft, review & editing.

All authors contributed to the literature review, design, data collection, drafting the manuscript, read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this study.

ETHICAL APPROVAL

According to review board and the ethics committee of Tehran University of medical sciences, retrospective analyses of fully anonymized hospital-based data are exempt from formal ethics approval. The study protocol was reviewed by the ethics committee of Tehran University of medical sciences, which confirmed that no formal approval was required. As only de-identified data were used, obtaining informed consent from individual patients was deemed unnecessary.

FINANCIAL DISCLOSURE

No financial interests related to the material of this manuscript have been declared.

AI USAGE

No AI was used for data analysis, interpretation, or content generation. The authors assume full responsibility for the final manuscript.

REFERENCES

1. Lavalley S, Caruso S, Foti R, Gagliano C, Cocuzza S, La Via L, et al. Behçet's disease, pathogenesis, clinical features, and treatment approaches: A comprehensive review. *Medicina (Kaunas)*. 2024; 60(4): 562. PMID: 38674208 DOI: 10.3390/medicina60040562 [[PubMed](#)]
2. Faza A, Makhlof Y, Ben Massoud F, Miladi S, Boussaa H, Ouenniche K, et al. Behçet disease: Epidemiology, classification criteria and treatment modalities. *Expert Rev Clin Immunol*. 2024; 20(12): 1437-48. PMID: 39101633 DOI: 10.1080/1744666X.2024.2388693 [[PubMed](#)]
3. Aljohani R, Alharbi S. Epidemiology and clinical features of Behçet's disease in the Middle East and North Africa: A systematic review and meta-analysis. *Clin Rheumatol*. 2026; 45(4): 2149-59. PMID: 41706278 DOI: 10.1007/s10067-026-07982-7 [[PubMed](#)]
4. Kutlug Agackiran S, Kardes E, Agackiran A, Direskeneli H, Alibaz-Öner F. Common femoral vein wall thickness as a supportive imaging marker in the differential diagnosis of Behçet's uveitis. *Turk J Med Sci*. 2026; 56(3): 761-8. PMID: 42499425 DOI: 10.55730/1300-0144.6209 [[PubMed](#)]
5. Bozkaya Yücel B, Doğanatan Ş, Çelik ÖG, Elmas S, Öğreten Yedigaroğlu N, Eser Öztürk H, et al. Clinical phenotype and long-term damage in paediatric Behçet's disease: Comparative performance of ISG, ICBBD and PEDBD classification criteria in a retrospective endemic cohort. *RMD Open*. 2026; 12(1): e006514. PMID: 41802833 DOI: 10.1136/rmdopen-2025-006514 [[PubMed](#)]
6. Abdullah HA, Ahmed SA. Frequency of HLA-B* 51 in patients with Behçet's disease and its relationship with the clinical findings in Kurdish population in Iraq. *Genetics and Molecular Research*. 2026; 25(8s): 1-8.
7. Vitale A, Caggiano V, Sbalchiero J, Lopalco G, Ragab G, Guerriero S, et al. The risk of developing ocular involvement among Behçet's disease patients presenting with mucocutaneous involvement at disease onset: Data from the international AIDA network Behçet's disease registry. *Ophthalmol Ther*. 2026; 15(1): 223-37. PMID: 41205142 DOI: 10.1007/s40123-025-01258-0 [[PubMed](#)]
8. Nokhatha SAA, Aljaberi N, Zeyoudi JA, Almarzooqi A, Soukieh F, Khan M. High prevalence of HLA-B51 and ocular involvement in Behçet's disease: A multicenter cross-sectional study. *Rheumatol Int*. 2025; 45(6): 140. PMID: 40349258 DOI: 10.1007/s00296-025-

- 05892-6 [[PubMed](#)]
9. Pei M, Qian Y, Liu X, Zhao C, Gao F, Zhang M. Visual prognosis of Behçet's uveitis patients in China and its associated factors. *Front Immunol.* 2026; 17: 1617653. PMID: 41924254 DOI: 10.3389/fimmu.2026.1617653 [[PubMed](#)]
 10. Demiroğlu H, Barışta İ, Dündar S. Risk factor assessment and prognosis of eye involvement in Behçet's disease in Turkey. *Ophthalmology.* 1997; 104(4): 701-5. PMID: 9111266 DOI: 10.1016/s0161-6420(97)30249-8 [[PubMed](#)]
 11. Alalwany H, Alganmi N, Bawazir Y, Mustafa M, Abusamra H, Banjar H, et al. Enhancing the diagnosis of Behçet's disease using machine learning: A comparative study on clinical data from Saudi Arabia. *Int J Telemed Appl.* 2026; 2026: 6157852. PMID: 41583997 DOI: 10.1155/ijta/6157852 [[PubMed](#)]
 12. Qi S, Jiang Y, Li H, Gong X, Li Q, Zhou X, et al. Detection of rheumatoid arthritis-associated interstitial lung disease: A systematic review and meta-analysis. *BMC Pulm Med.* 2026; 26(1): 336. PMID: 42324528 DOI: 10.1186/s12890-026-04417-6 [[PubMed](#)]
 13. Fu D, Tan L, Yang Q. Machine learning in systemic lupus erythematosus research: A bibliometric and visual analysis (2015–2025). *Egyptian Rheumatology and Rehabilitation.* 2026; 53(1): 53.
 14. Hammam N, Bakhiet A, El-Latif EA, El G, II, Samy N, Noor RAA, et al. Development of machine learning models for detection of vision threatening Behçet's disease (BD) using Egyptian college of rheumatology (ECR)-BD cohort. *BMC Med Inform Decis Mak.* 2023; 23(1): 37. PMID: 36803463 DOI: 10.1186/s12911-023-02130-6 [[PubMed](#)]
 15. Suwa A, Horita N, Ishido T, Takeuchi M, Kawagoe T, Shibuya E, et al. The ocular involvement did not accompany with the genital ulcer or the gastrointestinal symptoms at the early stage of Behçet's disease. *Mod Rheumatol.* 2019; 29(2): 357-62. PMID: 29583036 DOI: 10.1080/14397595.2018.1457424 [[PubMed](#)]
 16. Seyahi E. Phenotypes in Behçet's syndrome. *Intern Emerg Med.* 2019; 14(5): 677-89. PMID: 30747365 DOI: 10.1007/s11739-019-02046-y [[PubMed](#)]
 17. Sada K-E, Miyawaki Y, Yanai R, Kida T, Onishi A, Yoshimi R, et al. Development and validation of data-driven, decision tree-based algorithms for identifying Behçet's disease in claims data. *Int J Med Inform.* 2026; 209: 106266. PMID: 41506174 DOI: 10.1016/j.ijmedinf.2026.106266 [[PubMed](#)]
 18. Tangaro S, Lopalco G, Sabella D, Venerito V, Novielli P, Romano D, et al. Unraveling the microbiome-metabolome nexus: a comprehensive study protocol for personalized management of behçet's disease using explainable artificial intelligence. *Front Microbiol.* 2024; 15: 1341152. PMID: 38410386 DOI: 10.3389/fmicb.2024.1341152 [[PubMed](#)]
 19. Hatemi G, Seyahi E, Fresko I, Talarico R, Uçar D, Hamuryudan V. Behçet's syndrome: One year in review 2025. *Clin Exp Rheumatol.* 2025; 43(10): 1699-708. PMID: 41133359 DOI: 10.55563/clinexprheumatol/8xuim0 [[PubMed](#)]
 20. Sulu B, Karakoc A, Ak T, Dincer ZT, Senturk EF, Batur S, et al. Behçet's syndrome within the spectrum of systemic vasculitides. *Saudi J Ophthalmol.* 2025; 39(4): 311-22. PMID: 41367840 DOI: 10.4103/sjopt.sjopt.355.24 [[PubMed](#)]
 21. Parsaei A, Moradi S, Masoumi M, Davatchi F, Najafi A, Kooshki AM, et al. Predictive value of erythrocyte sedimentation rate and C-reactive protein in Behçet's disease activity and manifestations: A cross-sectional study. *BMC Rheumatol.* 2022; 6(1): 9. PMID: 35144674 DOI: 10.1186/s41927-021-00241-z [[PubMed](#)]