



The Role of Artificial Intelligence in Predicting Autoimmune Diseases: A Literature Review

Vihan Paswan

Research Supervisor
Soumita Sen

Abstract

Autoimmune diseases occur when the immune system mistakenly attacks the body's own healthy cells and tissues. Conditions such as rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), inflammatory bowel disease (IBD), and Type 1 diabetes can be difficult to diagnose because their symptoms may overlap with other diseases and may vary considerably between individuals. Delayed diagnosis can allow disease activity and organ damage to progress. Artificial intelligence (AI), particularly machine learning (ML), offers a potential solution by analysing large and complex datasets that are difficult for humans to process manually. This literature review examines how AI is being used for early disease prediction, diagnosis, biomarker discovery, disease classification, and treatment-response prediction. It considers techniques including Random Forest, Support Vector Machines, Logistic Regression, Ridge Regression, artificial neural networks, natural language processing, and deep learning. AI can integrate clinical records with genomic, transcriptomic, proteomic, metabolomic, and other biological information to support personalised medicine. However, significant limitations remain, including small datasets, poor data quality, overfitting, limited external validation, lack of interpretability, and patient data privacy concerns. Overall, existing evidence suggests that AI has considerable potential to transform autoimmune disease management from a reactive approach into a more predictive and personalised model. However, further validation is necessary before widespread clinical use.

Keywords: Artificial Intelligence, Machine Learning, Autoimmune Diseases, Rheumatoid Arthritis, Systemic Lupus Erythematosus, Predictive Medicine, Personalised Medicine, Multi-omics

Received 25 Sep., 2026; Revised 28 Sep., 2026; Accepted 05 Oct., 2026 © The author(s) 2026.

Published with open access at www.questjournals.org

I. Introduction

Autoimmune diseases are a group of complex disorders in which the immune system loses its ability to distinguish between foreign substances and the body's own tissues. Instead of protecting the body from pathogens, immune cells begin attacking healthy cells and organs. Autoimmune rheumatic diseases, including rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE), can cause persistent inflammation, pain, and progressive organ or tissue damage. The complexity of these conditions makes early and accurate diagnosis difficult.

Traditional diagnosis often involves several stages of clinical examination, blood testing, medical imaging, and exclusion of other diseases. A single test is rarely sufficient because autoimmune diseases can have overlapping symptoms and biological characteristics. For example, joint pain and inflammation may occur in both RA and SLE. Conventional biomarkers can also produce false-positive or false-negative results. Consequently, patients may experience delays before receiving an accurate diagnosis.

This challenge has increased interest in artificial intelligence (AI). AI refers to computational systems capable of performing tasks that normally require aspects of human intelligence. Machine learning (ML), a branch of AI, allows computers to identify patterns within data and use those patterns to make predictions. A systematic review of AI applications in autoimmune diseases found that Random Forest and Support Vector Machine models were among the most commonly used approaches, with applications including risk prediction, diagnosis, disease classification, progression, and treatment management.

This literature review examines how AI can contribute to predicting autoimmune diseases and supporting personalised treatment. It also evaluates the limitations and ethical challenges that must be addressed before AI can become a reliable part of routine medical practice.

II. Research Question

How can artificial intelligence and machine learning improve the prediction, diagnosis, and personalised treatment of autoimmune diseases, and what challenges limit their clinical application?

Hypothesis

AI-based models can improve the early prediction, diagnosis, and treatment-response assessment of autoimmune diseases by identifying complex patterns across clinical and biological data; however, their reliability depends on the quality, size, diversity, and transparency of the data used to train them.

III. Objectives

This review has four main objectives:

1. To understand how AI and ML are being applied to autoimmune diseases.
2. To examine how different AI techniques can support disease prediction and diagnosis.
3. To investigate the role of AI in personalised treatment and treatment-response prediction.
4. To identify the limitations, ethical concerns, and future opportunities associated with AI in autoimmune healthcare.

IV. Methodology

This paper uses a **literature-review methodology** rather than original clinical experimentation. This paper reviewed and organised existing research on artificial intelligence, machine learning, autoimmune diseases, and personalised medicine by major AI applications.

The literature considered includes research on rheumatoid arthritis, systemic lupus erythematosus, inflammatory bowel disease, Type 1 diabetes, and other autoimmune conditions. Particular attention was given to studies involving clinical records, blood biomarkers, genetic information, electronic medical records, multi-omics data, and treatment-response prediction.

The reviewed literature shows that AI research in autoimmune diseases has expanded substantially. A systematic review published in *npj Digital Medicine* identified 169 eligible studies from 702 records and found that RA, multiple sclerosis, and IBD were among the most frequently investigated diseases. More recent reviews demonstrate continued expansion into multimodal and multi-omics approaches.

V. Understanding Autoimmune Diseases and Personalised Medicine

The immune system normally protects the body against harmful organisms such as bacteria and viruses. In autoimmune disease, this protective mechanism becomes dysfunctional. T cells and B cells may fail to recognise the body's own tissues as "self," contributing to an inappropriate immune response.

One major challenge is **heterogeneity**. Patients with the same autoimmune disease may have different genetic characteristics, symptoms, disease progression, and responses to treatment. Therefore, a treatment that works well for one patient may be less effective for another.

This problem makes personalised medicine particularly important. Instead of providing every patient with the same treatment, personalised medicine uses an individual's biological and clinical characteristics to determine the most appropriate treatment.

However, personalised medicine requires analysing large quantities of information. A patient's genetic profile alone can contain enormous numbers of variables. When combined with blood biomarkers, medical history, imaging, and treatment records, the resulting dataset becomes extremely difficult for a human to analyse manually.

AI can help identify relationships among thousands of variables simultaneously. Current research suggests that AI can integrate clinical, genetic, imaging, and other forms of information to support precision medicine in rheumatology.

VI. Machine Learning and Deep Learning

Machine learning enables computers to learn patterns from data rather than relying entirely on explicitly programmed rules. Several ML techniques have been investigated in autoimmune diseases.

Random Forest combines multiple decision trees to make predictions. It is useful when datasets contain many variables and can be applied to classification and risk prediction.

Support Vector Machines (SVM) classify data by identifying boundaries between different groups. SVMs have been frequently used in autoimmune disease research and were among the most common methods identified in an earlier systematic review.

Logistic Regression can estimate the probability of an outcome, such as whether a patient is at increased risk of developing a disease.

Deep Learning (DL) is a more advanced form of machine learning based on artificial neural networks. These networks contain interconnected layers that process information and identify increasingly complex patterns. DL is particularly useful for high-dimensional datasets, including medical images and biological data.

However, greater complexity can reduce transparency. A doctor may be able to understand why a conventional statistical model reached a particular conclusion, while a complex neural network may provide a prediction without an easily understandable explanation. This is known as the "**black box**" problem.

For medical decisions, this matters because doctors need to understand and evaluate the reasoning behind an AI prediction before relying on it.

VII. Multi-Omics: Combining Different Types of Biological Data

One of the most promising applications of AI is **multi-omics analysis**. Instead of examining only one biological characteristic, multi-omics combines different layers of biological information.

These may include:

- **Genomics** – information contained in DNA
- **Transcriptomics** – information about gene expression
- **Proteomics** – proteins present in cells or biological samples
- **Metabolomics** – small molecules involved in metabolism
- **Microbiomics** – information about microorganisms, such as gut bacteria

Traditional medical testing often examines a limited number of biomarkers at a time. AI can analyse multiple datasets simultaneously and search for relationships that may not be visible through conventional approaches.

Recent research in SLE has investigated integrating genomics, transcriptomics, proteomics, and metabolomics to identify molecular subgroups and improve risk modelling and treatment-response prediction.

This approach could help researchers discover **new biomarkers** that provide better information about disease risk or progression. It could also help divide patients into biologically meaningful subgroups rather than treating everyone with the same diagnosis as identical.

VIII. Natural Language Processing and Electronic Medical Records

Medical information is not always stored in a structured format. Doctors frequently record patient histories in clinical notes containing abbreviations, different terminology, and free-text descriptions.

Natural Language Processing (NLP) is an area of AI that enables computers to analyse human language. In healthcare, NLP can extract useful information from electronic medical records (EMRs).

For autoimmune diseases, this may include previous diagnoses, symptoms, medications, family history, and other medical conditions. Combining this information with laboratory and biological data could improve risk prediction.

However, electronic medical records can contain inconsistent terminology and missing information. Therefore, the usefulness of NLP depends heavily on the quality and consistency of the underlying records.

IX. AI for Diagnosis and Disease Prediction

AI has several potential applications before and during diagnosis. It can identify patients who have an increased probability of developing an autoimmune condition, distinguish between diseases with similar symptoms, and potentially identify disease subtypes.

For example, RA and SLE can present with overlapping symptoms such as joint pain. AI models trained using biological markers can identify patterns that distinguish the two conditions.

The literature reviewed in the source material reports a Random Forest model that differentiated RA and SLE using blood-protein information with **92.4% accuracy and an AUC of 0.95**. These figures indicate strong performance within the study context. However, such results should not automatically be interpreted as proof that the model will perform equally well in every hospital or population.

More broadly, recent reviews report that machine-learning models have demonstrated promising performance in RA diagnosis and treatment prediction, while emphasising the need for larger datasets, external validation, and clinical testing.

X. AI for Predicting Treatment Response

An important advantage of predictive AI is its potential to determine whether a patient is likely to respond to a particular treatment.

Autoimmune diseases are often treated using several different medications. If the first treatment is unsuccessful, doctors may need to change the therapy. This process can take considerable time and may expose patients to ineffective treatments or unnecessary side effects.

AI could instead analyse a patient's biological and clinical characteristics and estimate the probability that a particular treatment will be successful.

The source literature reports an example in which **Ridge Regression** was used to analyse early genetic-expression information and predict response to Methotrexate six months later, achieving an AUC of approximately **0.78**. Another example involved artificial neural networks predicting early response to TNF inhibitors using baseline inflammatory markers, with an AUC of approximately **0.783**.

More recent research continues to investigate AI-based treatment prediction in RA and related diseases. A 2025 scoping review identified 89 studies examining AI for treatment-response prediction in RA and spondyloarthritis, while noting considerable variation in methodology and the continuing need for external validation.

These applications show how AI could move autoimmune medicine away from a trial-and-error approach toward more informed treatment selection.

XI. Measuring AI Performance

Medical AI models cannot be evaluated simply by asking whether they appear intelligent. Their predictions must be measured using statistical performance indicators.

One important measure is the **Area Under the Receiver Operating Characteristic Curve (AUC)**. An AUC closer to 1 generally indicates stronger ability to distinguish between different outcomes, while a value closer to 0.5 indicates performance similar to random classification.

Accuracy can also be useful, but it does not provide a complete picture. For example, a model may appear highly accurate on an unbalanced dataset but perform poorly at identifying a less common disease.

Researchers must therefore consider multiple measures and, importantly, test AI models on data that were not used during training.

XII. Challenges and Limitations

Despite its potential, AI cannot automatically solve the problems of autoimmune disease prediction.

12.1 Limited Data

Autoimmune diseases are complex and, in some cases, relatively uncommon. Rare diseases may have particularly small patient populations. If an AI model is trained on too few examples, it may fail to learn general patterns.

12.2 Overfitting

Overfitting occurs when a model learns the training dataset too closely instead of learning patterns that apply generally. Such a model may perform extremely well during training but poorly on new patients.

Earlier research demonstrated that independent testing and cross-validation were not consistently used in autoimmune AI research, highlighting the importance of stronger validation practices.

12.3 Data Quality

AI predictions are only as reliable as the data used to train the model. Medical records can contain missing values, inconsistent terminology, measurement errors, and coding differences.

12.4 Lack of Explainability

Doctors must be able to understand and critically evaluate AI recommendations. If a model cannot explain why it produced a particular prediction, clinicians may hesitate to use it.

12.5 Privacy and Cybersecurity

Autoimmune disease research may involve extremely sensitive information, including genetic data and medical histories. Poorly protected datasets could expose patients to privacy risks. Therefore, data security and responsible data governance must be central to AI development.

12.6 Generalisability

A model trained using patients from one hospital or country may not perform equally well in another population. Differences in genetics, demographics, healthcare systems, medical terminology, and treatment practices can affect performance.

Recent reviews continue to identify small sample sizes, limited population diversity, overfitting, interpretability, and insufficient prospective validation as major barriers to clinical implementation.

XIII. Ethical Considerations

AI in healthcare raises important ethical questions. First, patients should understand when AI contributes to decisions about their health. Second, responsibility must remain clear: AI should support clinical decision-making rather than replace qualified healthcare professionals.

Bias is another concern. If training data do not adequately represent different populations, the resulting model may perform better for some groups than others. This could unintentionally increase healthcare inequalities.

Privacy is equally important. Genetic and medical information is highly sensitive, and researchers must ensure that patient information is securely stored and used responsibly.

Finally, AI systems should be continuously monitored. A model that performs well during research may behave differently when exposed to new patients or changing medical practices.

XIV. Future Directions

The future of AI in autoimmune disease research is likely to involve greater integration of different forms of data. Combining EMRs, genetic information, imaging, laboratory results, and multi-omics data could provide a more complete picture of each patient.

Researchers are also increasingly investigating methods that make AI models more interpretable. Explainable AI could help clinicians understand which biological or clinical factors contributed most strongly to a prediction.

Another important development is the movement towards larger, multi-centre datasets. Recent research in SLE suggests that AI-driven multi-omics approaches may help identify biologically distinct patient groups and improve prediction of disease progression and therapeutic response. However, these systems still require prospective, multi-centre validation before routine clinical use.

The ultimate goal should not be to replace doctors with AI. Instead, AI should function as an advanced analytical assistant, allowing doctors to use large amounts of information more efficiently while retaining human judgement and responsibility.

XV. Discussion

The evidence reviewed supports the hypothesis that AI has significant potential to improve autoimmune disease prediction, diagnosis, and personalised treatment. Machine learning can analyse complex relationships among clinical and biological variables that would be difficult to evaluate manually.

The strongest potential lies in combining different forms of information. Multi-omics approaches can provide insights into the biological mechanisms of disease, while EMRs can provide information about a patient's medical history. Together, these data sources could create more comprehensive predictive models.

However, promising performance in research studies should not be confused with readiness for widespread clinical use. AI models must be externally validated, tested on diverse populations, evaluated prospectively, and assessed for clinical usefulness. Current literature continues to identify these requirements as important barriers to translation.

Therefore, AI should currently be viewed as a **decision-support technology rather than an autonomous medical authority**. Its greatest value may come from helping doctors recognise patterns, identify high-risk patients, discover biomarkers, and select potentially effective treatments.

XVI. Conclusion

Artificial intelligence is emerging as an important tool for predicting and managing autoimmune diseases. Conditions such as RA and SLE are difficult to diagnose and treat because they are biologically complex and vary considerably between individuals. Traditional approaches may therefore require lengthy investigations and repeated treatment adjustments.

AI and machine learning provide an alternative by processing large quantities of clinical and biological information. Algorithms such as Random Forest, Support Vector Machines, Logistic Regression, Ridge Regression, and neural networks have demonstrated potential for disease classification, risk prediction, biomarker discovery, and treatment-response prediction.

Integrating multi-omics data is particularly promising because it allows researchers to examine several layers of biology simultaneously. NLP can further expand AI capabilities by extracting information from electronic medical records.

Nevertheless, important challenges remain. Small datasets, poor data quality, overfitting, limited generalisability, lack of explainability, and privacy risks can reduce the reliability of AI systems. Consequently, strong validation and ethical safeguards are essential.

Overall, AI has the potential to shift autoimmune healthcare from a largely reactive model towards **earlier prediction, more accurate diagnosis, and personalised treatment**. Its future success, however, will depend

not only on developing more sophisticated algorithms but also on ensuring they are transparent, secure, scientifically validated, and used responsibly alongside medical expertise.

References

- [1]. Stafford, I. S., Kellermann, M., Mossotto, E., Beattie, R. M., MacArthur, B. D., Ennis, S., et al. (2020). *A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases*. npj Digital Medicine, 3, 30.
- [2]. Danieli, M. G., Brunetto, S., Gammeri, L., Palmeri, D., Claudi, I., Shoenfeld, Y., & Gangemi, S. (2024). *Machine learning application in autoimmune diseases: State of the art and future perspectives*. Autoimmunity Reviews, 23(2), 103496.
- [3]. Momtazmanesh, S., Nowroozi, A., & Rezaei, N. (2022). *Artificial Intelligence in Rheumatoid Arthritis: Current Status and Future Perspectives: A State-of-the-Art Review*. Rheumatology and Therapy, 9, 1249–1304.
- [4]. *Advancing precision rheumatology: applications of machine learning for rheumatoid arthritis management*—Frontiers in Immunology (2024).
- [5]. *Artificial Intelligence in Rheumatology: Clinical Applications in Rheumatoid Arthritis, Osteoarthritis, and Systemic Lupus Erythematosus*. Cureus (2025).
- [6]. *Artificial intelligence to predict treatment response in rheumatoid arthritis and spondyloarthritis: a scoping review*. Rheumatology International (2025).
- [7]. Biswas, B., Munquad, S., Roy Choudhury, K., Cherayil, B. J., & Chaudhuri, A. (2026). *Applications of artificial intelligence in systemic lupus erythematosus: integrating multi-omics data for precision medicine*. Frontiers in Immunology, 17, 1804598.