Personalized medicine in 2025

https://doi.org/10.1038/s41584-025-01337-3

Advances in AI-based patient stratification for rheumatic diseases

Rachel Knevel

Advances in artificial intelligence (AI) are transforming patient stratification in rheumatology. In 2025, three landmark studies demonstrated how multimodal AI approaches spanning clinical, molecular and longitudinal data can uncover distinct disease subtypes and predict therapeutic response, advancing the field towards precision rheumatology.

The past year has seen artificial intelligence (AI) move from proof of concept to practical frameworks for stratifying patients with rheumatic diseases. AI is defined as models that learn complex, often nonlinear patterns from data with limited human supervision, such as deep neural networks, transformers or large multimodal architectures. These systems perform inference or prediction beyond explicit programming. Building on decades of work in clinical clustering and molecular profiling, researchers are now integrating complex, multimodal data — from transcriptomes to electronic health records (EHRs) and longitudinal registries — to define coherent patient subgroups. Three studies published in 2025 exemplify this progress (Table 1), showcasing different types of AI methods and tasks, collectively marking a pivotal step towards precision rheumatology.

At the molecular scale, Lewis et al. leveraged synovial RNA sequencing from the STRAP precision medicine trial ($n = 208$ people with rheumatoid arthritis (RA)) to train machine learning models to predict the response to three classes of biologic drugs: etanercept, tocilizumab and rituximab[1]. The resulting algorithms showed good performance and were externally validated using the independent R4RA cohort, confirming robustness in performance.

Across treatments, those who responded to treatment showed coordinated B cell and T cell activation, whereas those who did not displayed fibroblast and matrix-remodelling gene signatures. B cell modules predicted a response to etanercept and rituximab, and myeloid and interferon modules predicted a response to tocilizumab. These signatures were distilled into a 524-gene NanoString nCounter panel for clinical patient stratification.

Within the STRAP cohort, the authors reproduced the classical Pathobiology of Early Arthritis Cohort (PEAC) pathotypes[2] — 'lympho-myeloid', 'diffuse-myeloid' and 'pauci-immune fibroid' — and linked them to the Accelerating Medicines Partnership cell-type abundance phenotype (CTAP) framework[3]. Shared cellular axes underpinned treatment-response signatures: enrichment for IL-1B$^+$ macrophages and B cell subsets predicted a response to etanercept, whereas NUPR1$^+$ and interferon-activated macrophages characterized a response tocilizumab. This bridge between bulk and single-cell frameworks shows how the PEAC and CTAP systems converge on shared cellular axes of inflammation that define molecular subtypes and therapeutic sensitivity.

At the clinical level, Maarseveen et al. applied multimodal deep learning to EHR data from 1,387 patients with early RA to identify latent patterns in joint involvement, serology and inflammation[4]. The model revealed four reproducible joint involvement patterns (JIPs) — 'foot-predominant', 'seropositive oligoarticular', 'hand-dominant' and 'polyarticular disease' — validated across several independent cohorts.

Whether patients belonged to one of the four JIPs was not explained by disease duration — the cluster with the longest duration had lower joint counts — which suggests that high-activity JIPs do not represent later disease stages. Instead, Maarseveen et al. describe distinct inflammatory configurations at presentation, defined by both activity and anatomical distribution[4].

JIPs showed distinct prognoses; the 'hand-dominant' subgroup (typically older adults who were seronegative) exhibited superior methotrexate retention and remission, whereas 'foot-predominant' and 'polyarticular disease' phenotypes had higher treatment failure rates. These associations were strongest among patients who were positive for anti-citrullinated protein antibodies. Histology from 194 biopsy-obtained synovial samples showed differences in inflammatory infiltrates and lining-layer hyperplasia: the 'seropositive oligoarticular' group showed low-grade synovitis; those with 'hand-dominant' JIPs had moderate inflammation but much higher stromal density; the 'foot-predominant' cluster displayed balanced moderate inflammation; and those with 'polyarticular disease' JIP exhibited the most severe inflammatory activity. These contrasts persisted after correction for disease activity.

A complementary commentary by Ospelt and Ciurea proposed that biomechanical load and joint-specific fibroblast programming influence the persistence of synovitis[5]. This perspective offers a plausible explanation for the poor prognosis of the 'foot-predominant' cluster, in which patients often experience treatment-resistant disease.

Key advances

- Molecular machine learning models trained on synovial RNA profiles predicted response to biologic therapies in rheumatoid arthritis (RA), enabling development of a streamlined 524-gene panel for clinical implementation[1].
- Deep learning in routine clinical data identified four reproducible RA phenotypes with distinct prognoses and histopathology, highlighting the importance of considering joint location for treatment decisions[4].
- Semi-supervised multi-organ modelling uncovered five hierarchical subtypes of systemic sclerosis, providing novel insights into patients with high-risk disease[6].

Table 1 | 2025 research on AI-based patient stratification in rheumatic diseases

Study	Study design	Patients (patient number)	Type of AI method	Primary findings
Lewis et al. (STRAP cohort)[1]	Prospective precision medicine trial with external validation using RNA-sequencing data from synovial samples	Patients with RA treated with etanercept, tocilizumab or rituximab (n = 208 for the STRAP cohort, and n = 100 for the R4RA validation cohort)	Supervised machine learning (predictive modelling and feature selection on transcriptomic data)	Machine-learning models predicted biologic therapy response (AUC 0.75–0.79). Those who responded to treatment showed B cell and T cell activation, and those who did not had fibroblast and ECM remodeling. A 524-gene NanoString panel reproduced the predictive signatures.
Maarseveen et al. (JIP framework)[4]	Multimodal deep learning analysis of EHRs	Early RA cohorts (n = 1,387 for discovery cohort, and n ≥ 2,000 validation cohort)	Unsupervised deep learning (representation learning and clustering)	Identified four reproducible JIPs (foot, hand, oligo and poly) with distinct prognoses and histopathology. Differences in lining hyperplasia and infiltrate persisted after correction for disease activity, defining stable clinical–histological subsets.
Trottet et al. (EUSTAR registry)[6]	Semi-supervised generative deep learning framework on longitudinal registry data	Patients with systemic sclerosis in the EUSTAR registry (n = 14,000)	Semi-supervised generative modelling (trajectory prediction across multi-organ features)	The model learned eight organ trajectories and defined five hierarchical multi-organ subtypes. The 'limited-skin but high-risk' subgroup was predictive of early lung and heart involvement (F1 = 0.77–0.83). Demonstrated robust multi-organ disease-trajectory modelling.

AI, artificial intelligence; AUC, area under the receiver operating curve; ECM, extracellular matrix; EHRs, electronic health records; EUSTAR, European Scleroderma Trials and Research; JIP, joint involvement patterns; RA, rheumatoid arthritis.

Taken together, these insights position the JIP framework as the clinical systems counterpart to molecular and cellular stratification efforts such as PEAC and Accelerating Medicines Partnership CTAP. By demonstrating that reproducible disease structure is detectable in routine EHR data, Maarseveen et al. extend precision rheumatology beyond tissue analysis[4].

Beyond RA and single-organ models, Trottet et al. applied a semi-supervised generative deep learning framework to 67,000 clinical visits from 14,000 patients in the European Scleroderma Trials and Research (EUSTAR) systemic sclerosis registry[6]. The model integrated longitudinal data across eight organ systems and expert severity scores to learn multi-organ disease trajectories.

The algorithm identified five hierarchical subtypes of systemic sclerosis spanning a mild-to-severe continuum. The 'limited-skin but high-risk' group showed elevated risk of lung and heart involvement despite minimal cutaneous disease. Other subtypes captured musculoskeletal or vascular trajectories. Modelling disease as a trajectory enabled prediction of organ involvement several visits in advance of the development of such involvement. Clinician-guided semi-supervision enhanced interpretability by aligning latent dimensions with recognizable organ systems. This work demonstrates how AI can synthesize incomplete, heterogeneous longitudinal data into coherent temporal phenotypes — an approach that parallels a trajectory-based approach using graph-based pseudotime clustering published in 2025 (ref. 7).

Although AI can reveal striking patterns, the value of this information depends on careful clinical and epidemiological interpretation. A central question in RA stratification is whether observed heterogeneity reflects fixed subsets or transient states[8]. Triaille et al. argued that correlations between molecular profiles and current disease activity favour a state-based interpretation[8]. However, the Maarseveen study shows that histological differences between RA clusters persist after adjustment for disease activity, and that clinical trajectories do not simply follow disease duration[4]. Longitudinal analyses — such as the trajectory modelling in systemic sclerosis by Trottet et al.[6], and similar trajectory work now emerging in RA[7] — provide further insight into whether differences are stable over time. Together, these findings suggest that activity-driven states alone cannot account for the diversity observed. A more coherent view is that RA heterogeneity arises from interactions between dynamic inflammatory states and more stable, anatomy or tissue-linked traits. This view reconciles the critique from Triaille et al.[8] with newer data; inflammation shapes much of the measurable variation, but it unfolds upon underlying predispositions that confer stability to specific disease patterns.

The three stratification studies highlighted here[1,4,6] collectively mark a transition from static classification to dynamic, data-driven stratification across rheumatic diseases. First, at the molecular level, transcriptomic models connect tissue biology with therapeutic response[1]. Second, at the clinical level, deep learning on routine data exposes structures that are invisible to standard scores[5]. Finally, at the systems level, generative modelling reconstructs multi-organ disease trajectories[6].

A unifying feature is the fusion of expert knowledge and AI, which enhances interpretability and provides important mechanistic context for AI discoveries. Rigorous replication[1,4] and expert-guided validation[6] reinforce confidence that these findings are biologically grounded. The next phase of disease stratification will probably adopt generative AI principles, including those seen in large language models exemplified by population-scale transformers such as DELPHI-2M, which learned long-term disease trajectories and could forecast risk decades ahead of time[9].

The advances in the three papers[1,4,6] emphasize that precision stratification depends on large, granular and diverse datasets. Future progress will rely on capturing all dimensions of rheumatic diseases, from detailed longitudinal EHRs and multi-omics tissue or blood data to imaging and, where feasible, routine biopsy-obtained synovial samples. Even at population scale, as illustrated by the DELPHI-2M model[9], the temporal sequence of diagnoses and clinical notes carries strong predictive power. For rheumatic diseases, richer data coupled with clinician-guided annotation can accelerate access to the correct therapy and inform strategies for early intervention or disease prevention.

Progress also requires inclusion of all patients. As emphasized in the 2025 *Nature Reviews Rheumatology* Editorial, "the optimal path forward is to ensure diversity in every aspect of clinical practice, so that the breakthrough of early diagnosis and equitable care can be celebrated at the journal's next notable anniversary."[10] Achieving true precision demands that data from each patient contribute to discovery. AI offers a route to this goal, as it enables stratification within heterogeneous populations rather than necessitating homogeneous cohorts, as in traditional trials. The studies highlighted here demonstrate that AI could enhance biological studies and clinical trials by defining biologically coherent, reproducible subsets of patients.

Rachel Knevel [1,2,3]
[1]Department of Rheumatology, Leiden University Medical Center, Leiden, the Netherlands. [2]Translational & Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK. [3]Delft Bioinformatics Lab, Delft University of Technology, Delft, the Netherlands.
e-mail: r.knevel@lumc.nl

Published online: 3 December 2025

References

1. Lewis, M. J. et al. STRAP collaborative group. Deep molecular profiling of synovial biopsies in the STRAP trial identifies signatures predictive of treatment response to biologic therapies in rheumatoid arthritis. *Nat. Commun.* **16**, 5374 (2025).
2. Lewis, M. J. et al. Molecular portraits of early rheumatoid arthritis identify clinical and treatment response phenotypes. *Cell Rep.* **28**, 2455–2470 (2019).
3. Zhang, F. et al. Deconstruction of rheumatoid arthritis synovium defines inflammatory subtypes. *Nature* **623**, 616–624 (2023).
4. Maarseveen, T. D. et al. Location and amount of joint involvement differentiates rheumatoid arthritis into different clinical subsets. *npj Digit. Med.* **8**, 623 (2025).
5. Ciurea, A. & Ospelt, C. Location-specific treatment of chronic inflammatory joint disease. *Nat. Rev. Rheumatol.* **21**, 507–508 (2025).
6. Trottet, C. et al. EUSTAR Collaborators. Deep hierarchical subtyping of multi-organ systemic sclerosis trajectories – a EUSTAR study. *npj Digit. Med.* **8**, 563 (2025).
7. Steinz, N. et al. Time-independent disease state identification defines distinct trajectories determined by localised vs systemic inflammation in patients with early rheumatoid arthritis. *Ann. Rheum. Dis.* https://doi.org/10.1016/j.ard.2025.04.011 (2025).
8. Triaille, C. et al. The relationship between clinical disease activity, synovial inflammatory profile, and treatment response in rheumatoid arthritis. *Lancet Rheumatol.* https://doi.org/10.1016/S2665-9913(25)00252-8 (2025).
9. Shmatko, A. et al. Learning the natural history of human disease with generative transformers. *Nature* **647**, 248–256 (2025).
10. The next breakthrough in rheumatology will require prioritizing diversity. *Nat. Rev. Rheumatol.* **21**, 641 (2025).

Competing interests

The author declare no competing interests.