

## Location and amount of joint involvement differentiates rheumatoid arthritis into different clinical subsets

Maarseveen, Tjardo D.; Maurits, Marc P.; Coletto, Lavinia Agra; Perniola, Simone; Böhringer, Stefan; Steinz, Nils; Reinders, Marcel J.T.; van den Akker, Erik B.; Knevel, Rachel; More Authors

**DOI**

[10.1038/s41746-025-01997-1](https://doi.org/10.1038/s41746-025-01997-1)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

npj Digital Medicine

**Citation (APA)**

Maarseveen, T. D., Maurits, M. P., Coletto, L. A., Perniola, S., Böhringer, S., Steinz, N., Reinders, M. J. T., van den Akker, E. B., Knevel, R., & More Authors (2025). Location and amount of joint involvement differentiates rheumatoid arthritis into different clinical subsets. *npj Digital Medicine*, 8(1), Article 623. <https://doi.org/10.1038/s41746-025-01997-1>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

<https://doi.org/10.1038/s41746-025-01997-1>

# Location and amount of joint involvement differentiates rheumatoid arthritis into different clinical subsets



Tjardo D. Maarseveen<sup>1</sup>, Marc P. Maurits<sup>1</sup>, Lavinia Agra Coletto<sup>2</sup>, Simone Perniola<sup>3</sup>, Stefan Böhringer<sup>4</sup>, Nils Steinz<sup>1</sup>, Sytske Anne Bergstra<sup>1</sup>, Dario Bruno<sup>3</sup>, Maria Rita Gigante<sup>3</sup>, Viviana A. Pacucci<sup>2</sup>, Luca Petricca<sup>2</sup>, Bianca Boxma-de Klerk<sup>1</sup>, Herman Kasper Glas<sup>5</sup>, Clara Di Mario<sup>6</sup>, Denise Campobasso<sup>6</sup>, Barbara Tolusso<sup>6</sup>, Josien Veris-van Dieren<sup>5</sup>, Annette H. M. van der Helm-van Mil<sup>1,7</sup>, Elisa Gremese<sup>3,8</sup>, Maria Antonietta D'Agostino<sup>2,6,8</sup>, Marcel J. T. Reinders<sup>9,10</sup>, Marco Gessi<sup>11</sup>, Tom W. J. Huizinga<sup>1</sup>, Stefano Alivernini<sup>2,6,8</sup>, Erik B. van den Akker<sup>9,10</sup> & Rachel Knevel<sup>1,12</sup> ✉

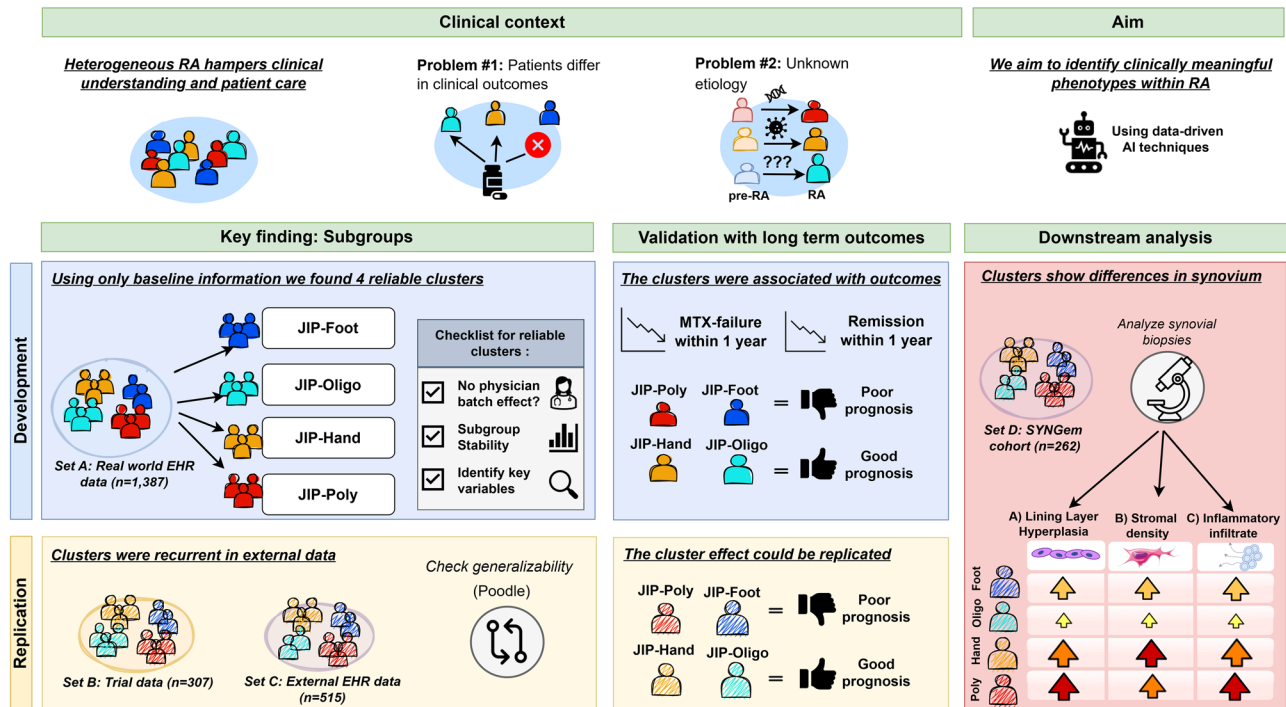
Rheumatoid arthritis (RA) is a heterogeneous disease with variable symptoms, prognosis, and treatment response, necessitating refined patient classification. We applied multimodal deep learning and clustering to identify distinct RA phenotypes using baseline clinical data from 1,387 patients in the Leiden Rheumatology clinic. Four Joint Involvement Patterns (JIP) emerged: foot-predominant arthritis, seropositive oligoarticular disease, seronegative hand arthritis, and polyarthritis. Findings were validated in clinical trial data ( $n = 307$ ) and an independent secondary care cohort ( $n = 515$ ). Clusters showed high stability and significant differences in remission rates ( $P = 0.007$ ) and methotrexate failure ( $P < 0.001$ ). JIP-hand patients had superior outcomes (particularly in ACPA-positive patients) versus JIP-foot (HR:0.37,  $P < 0.001$ ) and JIP-poly (HR:0.33,  $P = 0.005$ ), independent of baseline disease activity and clinical markers. Synovial histology analysis ( $n = 194$ ) revealed distinct inflammatory patterns across clusters, hinting at different underlying biological mechanisms. These validated RA phenotypes based on joint involvement patterns may enable targeted research into disease mechanisms and personalized treatment strategies.

Rheumatoid arthritis (RA) is a heterogeneous disease. The current classification criteria for RA were developed to approximate the decision to start early treatment and exclude other diseases. At clinical presentation, patients vary in the number and pattern of joints involved, presence of extra-articular manifestations, and abnormalities in blood and synovial tissue<sup>1–3</sup>. The heterogeneity of RA also manifests in clinical outcomes, namely prognosis, treatment response, and comorbidities. This evident diversity likely impacts the interpretation of treatment effects and etiologic factors such as genetics, and may downplay their importance altogether<sup>4</sup>. If phenotypic subsetting into more

homogeneous groups is possible, it could improve research into the etiology of RA and enhance its treatment.

For centuries, pattern recognition on clinical variables by doctors has been the driving force of disease identification and examination of the underlying etiologic mechanisms. Thus far, clinicians have not identified the relevant (sub)patterns in RA. The presence of ACPA<sup>5–8</sup> and the age of onset<sup>9,10</sup> have been raised as possible dichotomous disease subsetting features. However, neither of these markers in isolation adequately addresses the heterogeneity and complexity of the disease. This suggests that other factors are involved.

<sup>1</sup>Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands. <sup>2</sup>Division of Rheumatology – Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy. <sup>3</sup>Division of Clinical Immunology – Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy. <sup>4</sup>Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Zuid-Holland, Netherlands. <sup>5</sup>Rheumatology outpatient clinics, Reumazorg Zuid West Nederland, Goes, Netherlands. <sup>6</sup>Immunology Research Core Facility – Gemelli Science and Technology Park (GSTeP) – Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy. <sup>7</sup>Department of Rheumatology, Erasmus Medical Center, Rotterdam, Netherlands. <sup>8</sup>Università Cattolica del Sacro Cuore, Rome, Italy. <sup>9</sup>Leiden Computational Biology Centre, Leiden University Medical Center, Leiden, Zuid-Holland, Netherlands. <sup>10</sup>Delft Bioinformatics Lab, Delft University of Technology, Delft, Zuid-Holland, Netherlands. <sup>11</sup>Institute of Pathology – Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy. <sup>12</sup>Rheumatology, Newcastle University Translational and Clinical Research Institute, Newcastle upon Tyne, United Kingdom. ✉e-mail: [r.knevel@lumc.nl](mailto:r.knevel@lumc.nl)



**Fig. 1 | Workflow illustrating the study outline.** This study used data-driven analysis to clarify the complexity of rheumatoid arthritis, providing insights into disease etiology and predicting clinical outcomes. Cluster analysis identified four distinct baseline subgroups that differed in MTX treatment response and 1 year

remission rates. These clusters and treatment effects were replicated across multiple centers. Additional analysis of a separate cohort showed these clusters also corresponded to synovial differences. The fibroblast vector is from Servier Medical Art (<https://smart.servier.com/>), licensed under CC BY 4.0.

Cluster analysis combining a high number of factors has demonstrated its effectiveness in categorizing complex diseases (such as type II diabetes, asthma, and osteoarthritis) into subtypes that differ in clinical outcomes or biological background<sup>11–13</sup>. In the context of RA, there is considerable focus on molecular phenotyping, such as that done by Lewis et al.<sup>14</sup>, who discovered patterns in synovial tissue at baseline, with the lymphoid-myeloid pathotype being a predictor of poor outcome at disease onset<sup>15</sup>. Others used clinical and comorbidity information for clustering and identified four subsets, including one that exhibited a higher likelihood of biological DMARD initiation<sup>16</sup>. Likewise, Curtis et al.<sup>17</sup> used clinical variables, though not exclusively at baseline, and identified five clusters that differed in disease activity, RA duration, and type of comorbidities. These outcomes are typically highly influenced by treatment decisions and events that occur independently of the specific RA type. Furthermore, detailed clinical information such as the pattern of involved joints may be relevant for disease differentiation, as exemplified by psoriatic arthritis (PsA)<sup>18</sup>; yet none of the previous studies capitalize on this information for clustering.

Electronic Health Records (EHR) data provide a powerful asset for clustering as they encompass a wide variety of data modalities (laboratory values, clinical examinations, demographics) that each offer a unique perspective on the patient's condition. EHRs are collected during routine clinical care and thus resemble the true patient population more closely than study populations collected with a particular hypothesis in mind. The diversity of data types, however, poses a methodological challenge due to structural differences between the data modalities. The recent surge of deep learning tools<sup>19</sup> offers the possibility of combining different EHR layers into a patient representation by extracting the (hidden) factors that capture the most variation in the data. At present, many machine learning (ML) techniques exist to learn the relevant (clinical) patterns and encode patients accordingly. These embeddings can be used to detect patient subgroups, identify patterns, build predictive models, or assist in making disease classifications. The literature reports that clustering on top of these embeddings typically outperforms conventional techniques in cases of high-dimensional or complicated data<sup>20–22</sup>.

In this study, we aimed to dissect the clinical heterogeneity of RA by using the symptoms at initial presentation, before external factors such as treatment interfere. We hypothesize that the location of the involved joints and the inflammatory patterns observed in the blood play a role in subsetting RA, similar to their significance in distinguishing PsA from RA<sup>18</sup>. To achieve this, we make use of advanced data-driven techniques to identify and analyze disease-differentiating signatures based on initial clinical variables and determine whether they relate to clinical outcomes and histological synovial features.

## Results

### Patients

We retrieved 2691 RA patients for training set A of whom, 1387 were included in our study based on the availability of lab values and joint counts. For the replication, set B and set C had 364 and 1227 RA cases, of whom 307 and 515 had complete information (Supplementary Table 1, Supplementary Fig. 1). For the downstream analysis, we looked at biopsy data from 264 patients of set D, of whom 194 patients had synovial tissue taken from the same joint (Supplementary Table 2). A workflow diagram for the different phases of the study is shown in Fig. 1.

Each dataset captured a typical early RA population<sup>23,24</sup>. In comparison to set A, patients in the replication sets exhibited higher rates of seropositivity and had fewer tender joints. On average, set B patients were younger and set C patients had less inflammation, with a median ESR of 16. We used the phenotypic variables (see methods) of set A to construct the patient embedding. We visualized how much each type of EHR data contributed to our final embedding using a flameplot analysis (Supplementary Fig. 2).

### Four clusters separated by joint location, serology and blood values

The patient embedding showed four distinct clusters (Supplementary Fig. 3 & 4) which were not determined by any single clinical variable, as indicated by the wide dispersion of values. The primary variation across the clusters was their differences in affected joints, leading us to name them Joint

Involvement Patterns (JIP). Additionally, the clusters differed in levels of inflammation, age, and seropositivity (Supplementary Table 3, Supplementary Fig. 5):

- **Cluster 1: JIP-foot**  
moderate number of involved joints, particularly feet joints, younger patients, low leukocyte and thrombocyte levels.
- **Cluster 2: JIP-oligo**  
limited joint involvement and mostly seropositive patients.
- **Cluster 3: JIP-hand**  
elderly patients, symmetrical polyarthritis of hands, seronegative.
- **Cluster 4: JIP-polyarthritis**  
majority seronegative polyarthritis in hand and feet though with lower ESR.

The clusters were stable, with an average of >80% of patients grouping together in the same cluster over the 1000 iterations in the stability analysis (Supplementary Figs. 6 and 7). In fact, the stability was better in our combined multimodal approach than if we take each data type (numeric/categorical) separately (Supplementary Fig. 8). The clusters were not driven by treating physicians (Supplementary Fig. 9) and were generalizable across different validation sets (Table 1, Supplementary Fig. 10 & 11), showing similar joint involvement patterns (Fig. 2). Also, the identified patient clusters did not seem to represent different disease stages, as the cluster with the longest symptom duration had the lowest joint count and vice versa. There were differences in cluster prevalences between the validation sets (Supplementary Table 4 and 5).

**Validation on clinical outcomes beyond baseline.** In set A, 80% of patients received MTX as an initial drug across all clusters. The Kaplan-Meier curves show a difference in MTX failure between the clusters: 27%, 23%, 16%, 30% (for clusters 1–4,  $P < 0.001$ , Fig. 3a). The JIP-hand had clearly the best prognosis, where patients were twice as likely to stay on MTX compared to the most severe disease subtype, JIP-poly (HR: 0.48 (95% CI 0.35–0.77),  $P < 0.001$ ). Additionally, the JIP-hand performed better than the JIP-foot (HR: 0.55 (0.37–0.82),  $P = 0.003$ ).

Consistent with MTX response, we observed differences in remission rates: 44.3%, 47.4%, 55.7%, 38.5% (for clusters 1–4,  $P = 0.007$ , Fig. 3b), with the biggest difference between the JIP-hand and JIP-poly (HR: 1.65 (95% CI 1.2–2.29),  $P = 0.002$ ), even when corrected for baseline disease activity (Supplementary Fig. 12).

**ACPA within the clusters.** Since the literature reports that ACPA is indicative of persistent disease<sup>25</sup>, we examined whether the ACPA status was the main factor driving the difference in MTX failure. In set A, we found a higher treatment failure in ACPA-positive than ACPA-negative patients (27.6% versus 22.0%, Fig. 4), though it was not significant ( $P = 0.057$ ). Moreover, the association of ACPA with MTX failure differed within the clusters ( $P < 0.001$ , Fig. 4).

The difference between JIP-hand and the foot clusters JIP-poly and JIP-foot was larger within the ACPA-positive stratum (JIP-hand versus JIP-foot: HR 0.37 (0.15–0.60),  $P < 0.001$ ; JIP-hand vs JIP-poly: HR 0.33 (0.15–0.72),  $P = 0.005$ ) (Fig. 4b). For remission, we could not find this difference in the ACPA-positive stratum.

The good response in JIP-hand raised the question of whether this group overrepresented patients with parvovirus-induced arthritis instead of RA, but none of our clusters were enriched for parvovirus-positive patients (Supplementary Fig. 13).

**Replication.** In set C, 79% of patients received MTX as the initial drug across all clusters, whereas in set B, all patients were administered MTX. In both

replication sets, we again observed superior outcomes for the JIP-hand cluster regarding MTX failure (global  $P < 0.001$  and  $P < 0.001$ ). Remission could only be tested in replication set B, where it confirmed our previous finding ( $P < 0.001$ ). Consistent with the original finding, the difference between JIP-hand and JIP-poly was particularly pronounced within the ACPA-positive stratum in both replication set B (OR: 0.22 (0.12–0.63),  $P = 0.017$ ) and set C (HR: 0.38 (0.22–0.68),  $P < 0.001$ ). Within the ACPA-positive stratum, we also found a significant difference between JIP-hand and JIP-foot in replication set C (HR: 0.37 (0.15–0.93),  $P = 0.034$ ), though not in replication set B (OR: 1.03 (0.34–3.05),  $P = 0.801$ ). All analyses remained significant when corrected for baseline DAS (Supplementary Fig. 14). Even after adjusting for symptom duration in sets A and B, the association between cluster and treatment persisted (Supplementary Fig. 15).

**Informative value of clusters beyond known risk factors for MTX failure.** To ensure that the association between clusters and treatment outcomes was not solely driven by established clinical markers, we adjusted the Cox regression model accordingly (Supplementary Fig. 16). Notably, the inclusion of well-known contributing factors such as ACPA, RF, and the number of affected joints did not reduce the additional predictive value of clustering ( $P = 0.020$  in Set A;  $P = 0.019$  in Set C).

**Diversity in synovial characteristics.** Next, we examined histological variations between patient clusters in set D according to the Krenn synovitis components<sup>26</sup>. Biopsies were taken from the most inflamed joint, typically the knee or wrist. Due to the limited number of samples from other joints, a thorough comparison across different anatomical locations was not feasible. Therefore, we primarily focused on a single, consistent location that was involved in most clusters—specifically, the knee joints.

In total, we analyzed synovial tissue from 194 patients, distributed across four clusters: 27 in JIP-foot, 49 in JIP-oligo, 86 in JIP-hand, and 32 in JIP-poly (see Supplementary Fig. 17). The Krenn synovitis score differed significantly among the clusters, with mean scores of 5, 4, 5, and 6, respectively ( $P = 0.004$ ). The highest synovitis score was observed in the JIP-poly cluster, while the lowest was seen in the JIP-oligo cluster (Fig. 5).

The clusters showed significant histopathological variations between the JIP variants in lining layer hyperplasia ( $P = 0.026$ ), stromal density ( $P = 0.045$ ), and inflammatory infiltrate ( $P = 0.002$ ). These variations were primarily driven by the distinction between the more stereotypical RA clusters (JIP-hand/JIP-poly) and JIP-oligo, which was characterized by low-grade synovitis. In JIP-poly, 43.8% had severe synovial lining hyperplasia compared to 14.3% in JIP-oligo ( $P = 0.002$ ). Moreover, 21.9% had severe inflammatory cell infiltration compared to 6.1% in JIP-oligo ( $P = 0.001$ ) and 18.8% had severe stromal density versus 8.2% for JIP-oligo ( $P = 0.034$ ). JIP-hand exhibited increased synovitis levels compared to JIP-oligo for lining layer hyperplasia (27.9% vs. 14.3%;  $P = 0.015$ ) and inflammatory infiltrate (17.4% vs. 6.1%;  $P = 0.013$ ). Furthermore, in JIP-hand, 29.1% had severe stromal density compared to 7.4% in JIP-foot ( $P = 0.211$ ) and 8.2% in JIP-oligo ( $P = 0.013$ ). Although JIP-foot had a similar overall synovitis score as JIP-hand, it was not characterized by any particular KSS component; rather, it had moderate inflammatory activity across all elements.

Since KSS was previously shown to be linked to disease activity<sup>3</sup>, we adjusted for DAS categories to ensure the cluster differences were not merely an effect of the clinical metric. After correction, differences in lining layer hyperplasia ( $P = 0.035$ ) and inflammatory infiltrate ( $P = 0.005$ ) persisted between clusters, while stromal density differences became non-significant ( $P = 0.082$ ).

For a complete overview, we also examined whether sampling from joint areas specific to each JIP phenotype would yield different results compared to using only knee biopsies. We collected samples from lower extremities for JIP-foot patients, knees for JIP-oligo patients, fingers or wrists for JIP-hand patients, and knees or wrists for JIP-poly patients. This targeted sampling approach produced results similar to those found when analyzing knee biopsies alone (Supplementary Fig. 18).

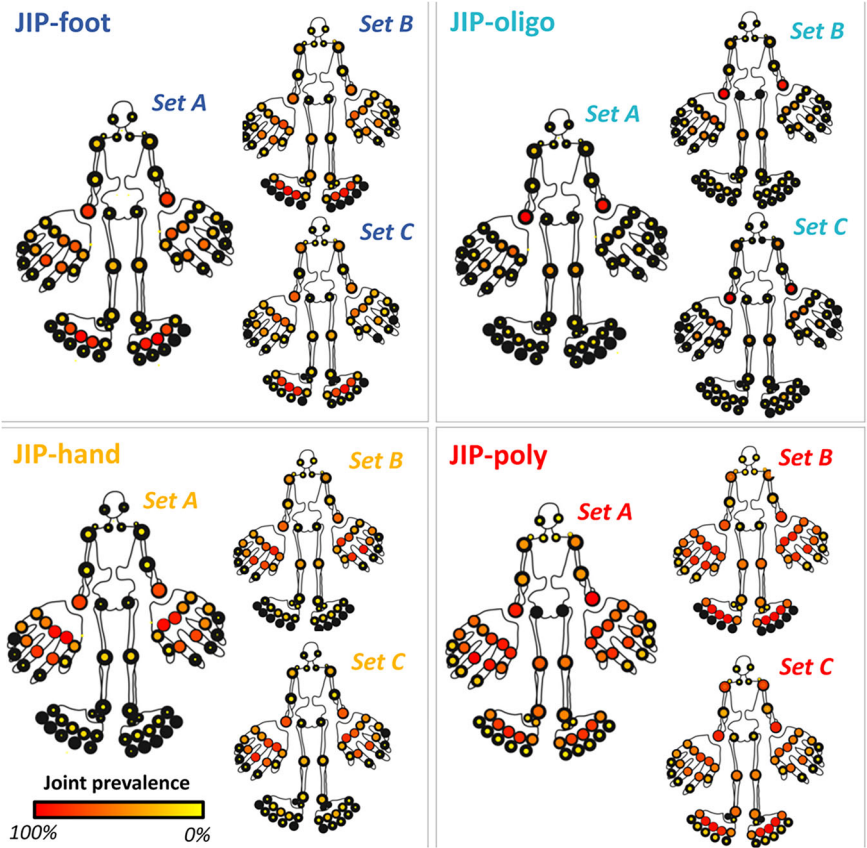


**Table 1 | Baseline characteristics of the different patient clusters (set A + replication sets B & C)**

	JIP-Foot	JIP-Oligo	JIP-Hand	JIP-Poly
N	596	761	450	402
Sex, female <sup>γ</sup> [n(%)]	389 (65.3)	505 (66.4)	272 (60.4)	272 (67.7)
Age <sup>γ</sup> (SD, yr)	56.6 (14.6)	59.5 (14.6)	66.4 (13.2)	54.7 (14.7)
RF <sup>γ</sup> [n(%)]	370 (62.1)	497 (65.3)	200 (44.4)	196 (48.8)
ACPA <sup>γ</sup> [n(%)]	355 (59.6)	449 (59.0)	162 (36.0)	183 (45.5)
ESR <sup>γ</sup> (IQR, mm/hr)	22 (9–36)	24 (11–38)	28 (13–48)	22 (9–40)
DAS44(3) (IQR)	3.5 (3.0–4.0)	2.4 (1.9–2.8)	3.7 (3.2–4.3)	4.6 (4.0–5.5)
SJC (IQR)	8 (5–12)	2 (1–5)	10 (7–15)	15 (9–22)
TJC (IQR)	11 (8–15)	3 (2–5)	11 (7–15)	22 (15–30)
DAS28(3) (IQR)	5.2 (4.5–6.0)	4.0 (3.3–4.7)	5.5 (4.9–6.2)	6.5 (5.5–7.2)
Follow up (IQR, days)	1307 (733–2020)	1428 (760–2082)	1127 (566–1875)	1512 (1012–2246)
Symptom duration * (IQR, days)	143 (56–364)	186 (70–399)	101 (48–279)	147 (56–357)

Where *SD* standard deviation, *RF* rheumatoid factor, *ACPA* anti-cyclic citrullinated peptide antibodies, *ESR* erythrocyte sedimentation rate, *IQR* interquartile range, *DAS* three component disease activity score (either 44 or 28 joint scheme), *SJC* swollen joint count, *TJC* tender joint count, <sup>γ</sup> clinical variables that were used for clustering (set A), \*symptom duration was only calculated for patients from set A and B.

**Fig. 2 | Pictorial mannequins for replication set B and C and their original counterpart (set A) to show the affected joints for each cluster with color and size to depict prevalence.** Frequency is colored on a gradient from red (=100%) to yellow (=0%). When there is no colored dot, it signifies the absence of both swelling and pain at baseline for these patients.

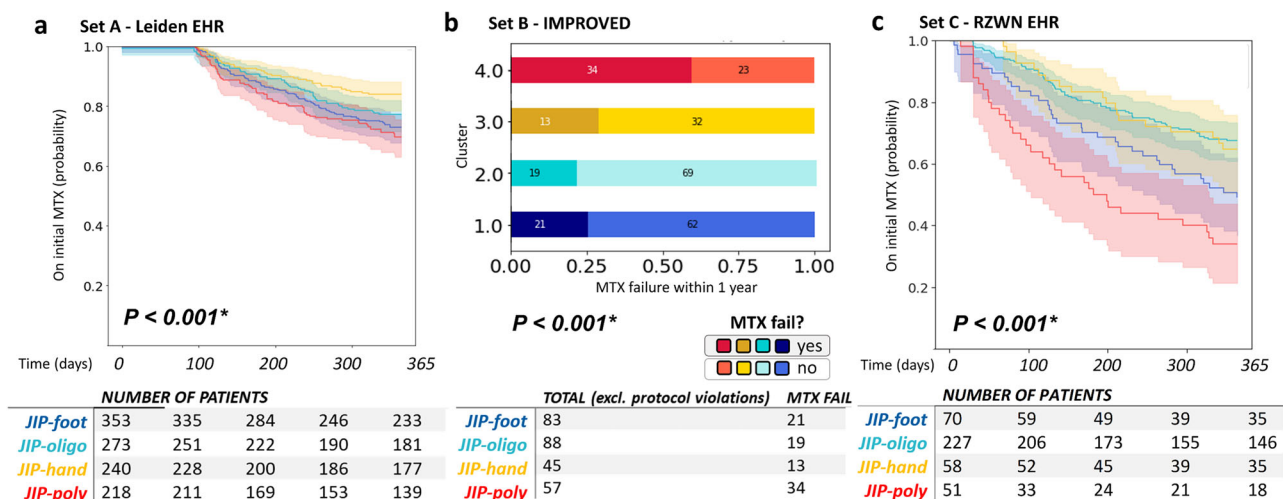


**Discussion**

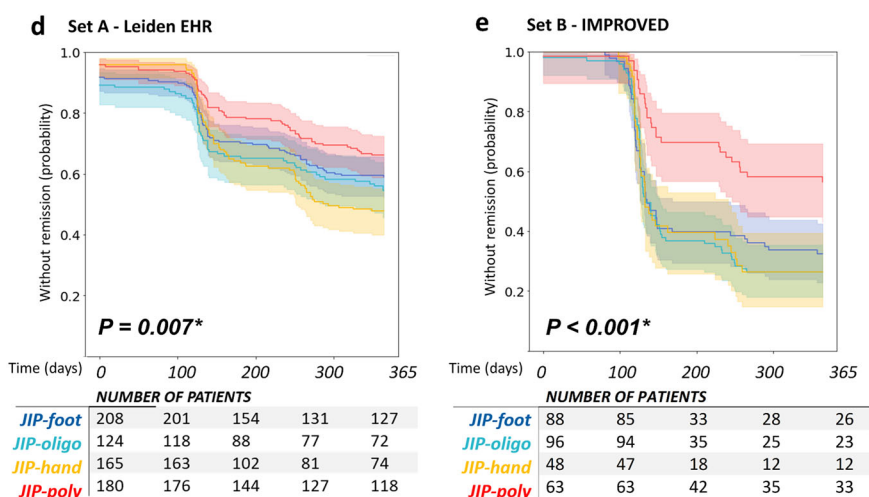
Through deep learning and clustering analysis of real-world clinical data, we identified four distinct RA phenotypes at baseline: foot-dominant (JIP-foot), seropositive oligoarticular (JIP-oligo), seronegative hand-dominant (JIP-hand), and polyarticular (JIP-poly) disease. While our hypothesis-free approach enabled detection of novel non-linear clinical signatures, we mitigated the risk of spurious correlations through rigorous validation, including stability testing, clinical outcome validation, independent cohort replication, and synovial histological correlation.

These results show promising directions for advancing RA management. While not yet clinically applicable, the identified clusters show clinical value as prognostic markers, with hand-foot differentiation proving particularly valuable for predicting clinical outcomes. Both foot-dominant clusters (JIP-foot, JIP-poly) showed higher MTX failure and lower remission rates compared to JIP-hand, independent of baseline joint involvement, symptom duration, or treatment timing. Notably, the impact of foot involvement on treatment failure rivaled that of ACPA-positivity as an independent risk factor,

## Methotrexate Failure



## Remission



**Fig. 3 | Association of baseline joint involvement clusters with clinical outcomes at 1 year. a** Kaplan–Meier curves for MTX switch rates in set A MTX-starters ( $n = 1,084$ ). **b** MTX switch rates in set B ( $n = 273$ ). **c** Kaplan–Meier curves for MTX switch rates in set C ( $n = 406$ ). **d** Remission rates (DAS44 < 1.6) in set A ( $n = 677$ ). **e** Remission rates in set B ( $n = 295$ ). Survival curves were generated for time-to-event

data (sets A and C), while cross-tabulations were used for set B, where MTX switching occurred at a fixed time per study protocol. Statistical comparisons were performed using log-rank tests for survival curves and chi-squared tests for cross-tabulated data.

indicating that a more comprehensive joint assessment should be considered in clinical practice.

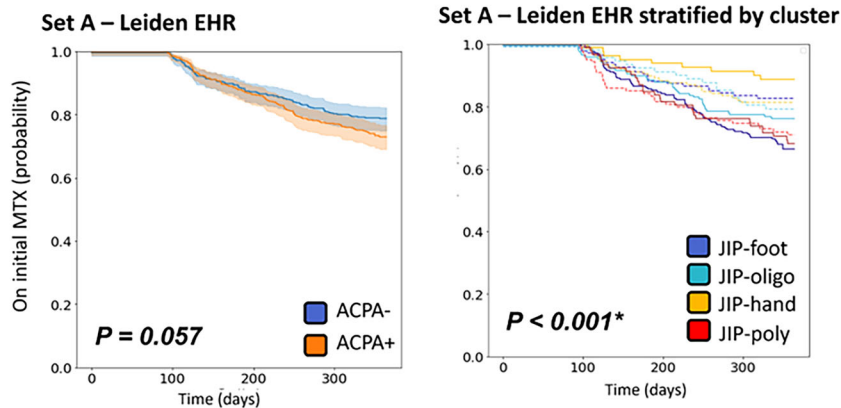
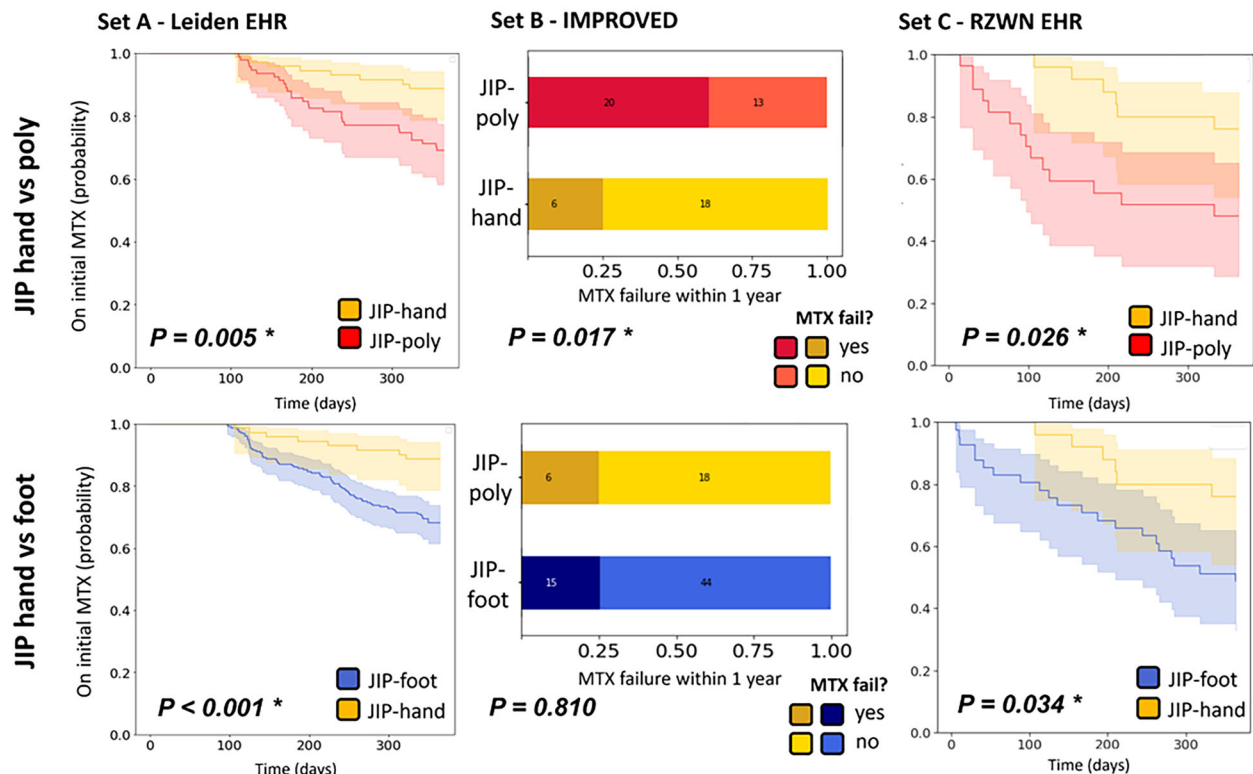
While this research underscores the importance of foot joints, current joint-specific treatment research often relies on 28-joint counts<sup>27</sup>, overlooking the feet despite their common involvement in RA<sup>28–30</sup>. Notably, our findings validate previous cross-sectional observations of poor prognosis in feet/ankle-involved disease<sup>28</sup> and support Ciurea et al.'s observation that foot involvement appears more persistent compared to other anatomical regions<sup>31</sup>. However, this study uniquely demonstrates this association in treatment-naïve patients at clinical presentation.

Nevertheless, dedicated trials are needed to determine whether tailoring therapy to specific patient groups—such as those with foot involvement—can improve outcomes compared to standard care. Recent studies demonstrate that synovial fibroblasts exhibit distinct transcriptomic and epigenomic profiles based on anatomical location, potentially explaining the biological underpinnings responsible for these joint-specific treatment outcomes<sup>32–34</sup>.

Contrary to common assumptions, we did not observe a clear ACPA dichotomy. This finding aligns with previous baseline studies that demonstrate that, despite known differences in risk factors and prognosis

between ACPA-positive and ACPA-negative patients, the clinical phenotype at initial diagnosis is similar for both groups<sup>25,35,36</sup>. ACPA prevalence was lowest in typical RA clusters (JIP-hand, JIP-poly) and highest in the oligoarticular cluster (JIP-oligo). While this pattern might partially reflect classification criteria<sup>37</sup>, our use of 1 year diagnosis validation and physician diagnosis in sets A and C minimizes misclassification bias. The high ACPA positivity in JIP-foot corroborates recent findings of increased foot involvement in ACPA-positive patients<sup>38</sup>.

Although seronegative patients have traditionally been considered to have a milder disease course<sup>39</sup>, our study shows this is not the case for the JIP-poly cluster, which is associated with poorer outcomes. In contrast, the seronegative JIP-hand cluster aligns with the expected pattern, showing a milder disease course. Recent literature also challenges the historical assumption of a uniformly mild prognosis in seronegative patients. For example, Duong et al. identified ACPA-positivity as a predictor of better treatment response<sup>40</sup>, and the ARCTIC trial by Haavardsholm et al. found that ACPA-negative patients exhibited slower treatment responses<sup>41</sup>. While our findings suggest that ACPA status may still play a role in disease progression, the relationship appears to be more nuanced and warrants further investigation.

**a General ACPA effects****b Within ACPA positive stratum**

**Fig. 4 | Methotrexate switch rates compared across ACPA status and joint involvement patterns within the seropositive stratum.** **a** Kaplan–Meier curves showing MTX switch rates by ACPA status overall and stratified by cluster, where dashed lines represent ACPA-positive patients. **b** Comparison of MTX switch rates

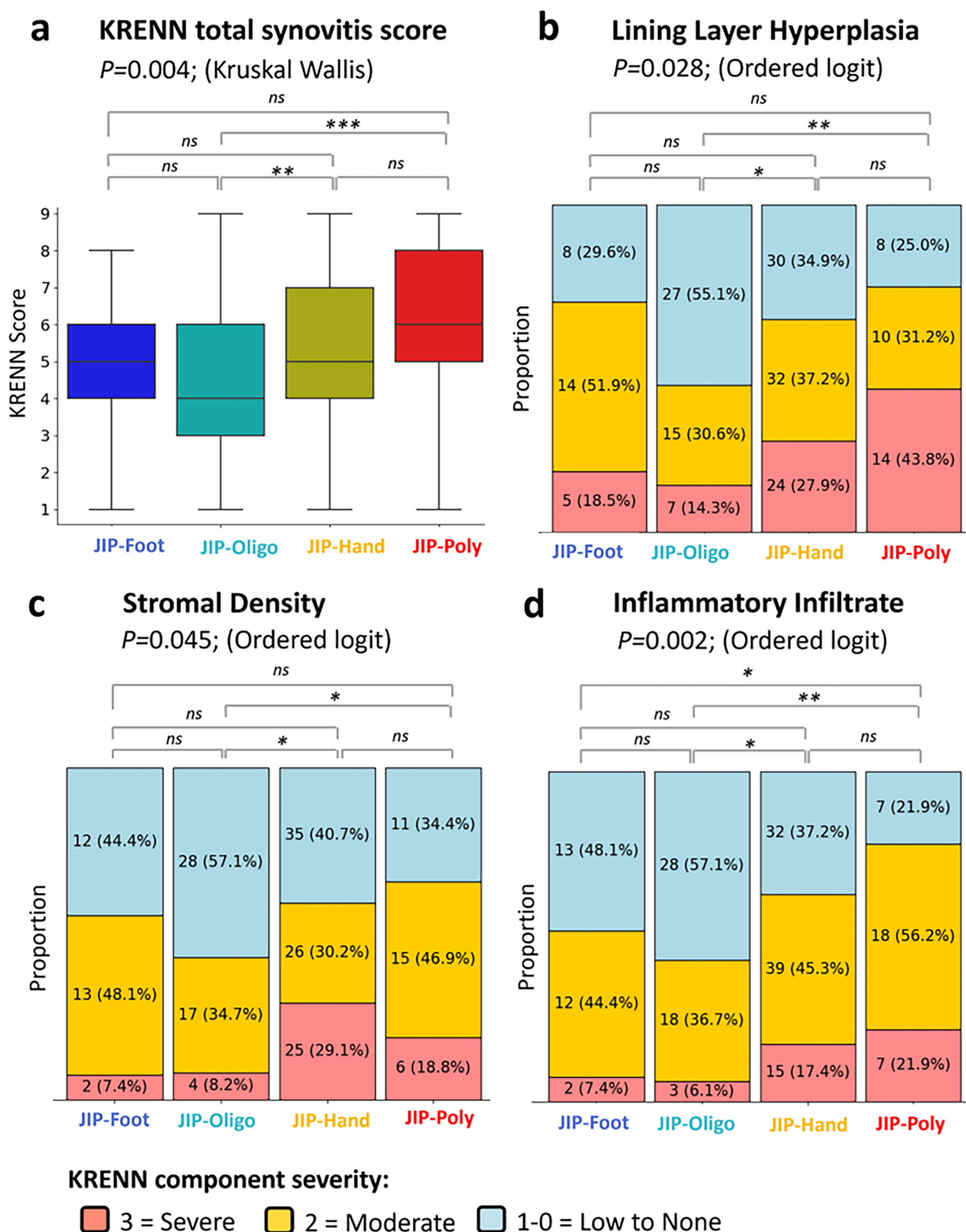
between JIP-hand and foot clusters (JIP-foot and JIP-poly) within ACPA-positive patients. Statistical comparisons used log-rank tests for survival curves and chi-squared tests for categorical data.

The treatment response disparity between hand and foot clusters was most pronounced in ACPA-positive patients. Across replication sets, we consistently observed significant differences between JIP-hand and JIP-poly, with two of three datasets also showing increased response in JIP-hand versus JIP-foot. For remission, we found that the different cluster-associated outcomes vanished in ACPA-positive patients, possibly reflecting the impact of targeted therapy intensification protocols particularly for ACPA positive patients.

Patients with hand-dominant joint inflammation (JIP-hand) showed surprisingly good outcomes, prompting us to explore several possible explanations. We found that these positive results could not be explained solely by how long patients had symptoms, their disease activity at baseline, or their lower prevalence of ACPA in that cluster. Additionally, we found no increased parvovirus positivity in the JIP-

hand group compared to other clusters, indicating that misdiagnosis of RA due to self-limiting reactive arthritis was unlikely<sup>42,43</sup>. One possible explanation is that these patients represent a phenotype of seronegative RA characterized by particular hand and wrist involvement, as reported by Burns et al.<sup>44</sup>. Alternatively, hand-dominant presentations may appear more responsive due to the widespread use of simplified disease indexes in trials (DAS28, CDAI) excluding foot and ankle joints<sup>28,30</sup>, potentially limiting the generalizability to atypical presentations.

Our clusters captured previously described age-related subsets, including elderly-onset RA (EORA) in JIP-hand, characterized by higher inflammation markers, lower female prevalence, and reduced autoantibody positivity<sup>10,11</sup>. However, our analysis revealed more granular subtypes beyond the EORA/YORA dichotomy.



**Fig. 5 | Synovitis severity across the patient clusters.** **a** Total Krenn synovitis scores for each cluster. **b** Lining layer hyperplasia scores by cluster. **c** Stromal density scores by cluster. **d** Inflammatory infiltrate scores by cluster. Global trends were assessed

using the Kruskal-Wallis test for total Krenn scores and ordered logit models for individual component grades, followed by post hoc Wald tests. Significance levels:  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ .



Further analysis of synovial tissue samples revealed distinct histological differences. Both the JIP-poly and JIP-hand groups exhibited severe synovitis. Specifically, JIP-poly was characterized by increased lining hyperplasia and sublining leukocytic infiltration, whereas JIP-hand showed greater stromal density. In contrast, the JIP-oligo group predominantly displayed mild inflammation. These histopathological differences were statistically significant. Notably, after adjusting for disease activity levels, the difference in stromal density was no longer significant, while the differences in inflammatory infiltrate and lining hyperplasia remained significant.

The striking histological differences between our patient clusters further supports the notion that these may represent distinct disease subtypes. The aggressive inflammatory patterns observed in JIP-poly and JIP-hand groups contrast sharply with the mild inflammation seen in JIP-oligo, suggesting different pathological mechanisms.

However, to fully understand the biological basis of these differences, research must extend beyond histopathological features alone. Molecular profiling of synovial tissue offers the potential for deeper mechanistic insights<sup>14</sup>. This approach has proven successful in previous work by Rivellese et al.<sup>45</sup> who identified distinct molecular pathotypes that predicted differential responses to tocilizumab and rituximab.

A limitation of our study is that we defined MTX success based on changes in medication, which could include switches due to side effects, though this likely underestimates rather than overestimates the observed associations. Center-specific therapeutic approaches varied but did not affect cluster-outcome associations. While temporal cluster stability was not directly assessed, previous evidence of consistent joint involvement patterns and cross-sectional associations support stability over time<sup>46</sup>.

Due to the real-world nature of the data, we could not account for all possible factors of influence. Most notably, missing BMI data may have limited insights into disease heterogeneity. Overweight RA patients show higher disease activity scores and worse treatment outcomes but less radiographic damage<sup>47,48</sup>, suggesting adiposity-related inflammatory pathways and cytokine profiles<sup>47–50</sup>. However, obesity may also inflate subjective clinical measures independent of actual inflammatory activity<sup>48,51</sup>, warranting future investigation.

Another limitation is that we solely focused on knee biopsies due to insufficient data for other joint locations. This prevented us from exploring and comparing different tissue environments even though they might be crucial for understanding the different phenotypes. Additionally, knee biopsies may reflect synovial features from concurrent non-inflammatory conditions such as osteoarthritis, which commonly affects this joint. However, a previous study by Alivernini et al.<sup>3</sup> demonstrates that Krenn synovitis scores are significantly lower in osteoarthritis biopsies ( $1.70 \pm 0.15$ ), suggesting minimal confounding impact on our assessment of inflammation.

Important to underline is that our identified clusters are not set in stone. Though we observed a high robustness of our clusters, patients lie on a gradient (Supplementary Fig. 4) and did not segregate in clearly separable modules. The cluster structure that we identified could also be summarized into more or fewer clusters and the clusters might become clearer when more layers of information are added. Such types of information could be genetics, gene expression patterns and molecular profiles from blood<sup>14,52</sup>. Despite these limitations, our study demonstrates the value of unsupervised, data-driven approaches in uncovering hidden disease patterns, with joint involvement patterns emerging as a major axis of variation.

In conclusion, our clustering analysis identified four baseline RA phenotypes, each defined by distinct patterns of hand and foot involvement that predict 1 year clinical outcomes and correspond to histological differences. This data-driven approach offers greater granularity than conventional age- or ACPA-based stratifications, suggesting the existence of distinct underlying etiologies that merit further biological investigation.

## Methods

### Patients

Our study comprises three different phases: a (i) developmental phase where we identify and validate subtypes in a discovery set according to long term

outcomes (set A), a (ii) replication phase where we cluster novel patients using historic trial- (set B) and external hospital data (set C) to infer generalizability by replicating the treatment analysis and finally (iii) a downstream analysis in external hospital data (set D) where we explore differences between clusters in their synovial tissue.

Set A consisted of 1387 RA-patients that visited the rheumatology outpatient clinic of the Leiden University Medical Centre for the first time between August 29th, 2011 till December 1st 2022. RA diagnosis was based on the physician's diagnosis within 1 year since first visit<sup>53,54</sup>.

Set B concerned 307 RA-patients from the IMPROVED trial that were recruited between March 2007 till September 2010<sup>55</sup>. This trial recruited undifferentiated arthritis and early RA with <2 years of symptoms. We selected only those patients who met the ACR2010 criteria within 1 year after inclusion. All patients received MTX at baseline and were randomized into two arms of treatment intensification if they did not reach remission after 4 months.

Set C included 515 RA patients from Reumazorg Zuid West Nederland (RZWN), collected between January 2015 and December 1, 2022. These patients were from nine different hospitals across the south west of the Netherlands, with the largest groups coming from Goes ( $n = 157$ ), Roosendaal ( $n = 153$ ), and Vlissingen ( $n = 49$ ). Herein, the diagnosis of RA was defined as having an ICD-code for RA and starting with a conventional DMARD.

Set D included 262 RA patients who met the ACR 2010 criteria for RA, drawn from the SYNGem Biopsy Unit cohort at the Fondazione Policlinico Universitario A. Gemelli IRCCS-Università Cattolica del Sacro Cuore in Rome, Italy. All patients underwent minimally invasive, ultrasound-guided synovial tissue biopsy during their initial rheumatological evaluation as part of routine clinical management<sup>3</sup>. Each tissue sample was processed with Hematoxylin and Eosin staining, and synovitis was graded using the total Krenn Synovitis Score (KSS)<sup>26</sup>.

Across all sets, a minimum follow-up of 1 year was required to ascertain the diagnosis of RA. Prior to conducting the study, we received ethical approval from the Medical Ethics Committee (METC) at Leiden University Medical Center according to study protocol B18.057. The study complied with the Declaration of Helsinki and national guidelines, including the COREON Code of Conduct for Health Research. Patients and the public were not involved in study development, execution, or dissemination.

### Preprocessing of electronic health records

To construct patient phenotypic profiles, we extracted information on serology (RF and ACPA), location of joint involvement (tender- and swollen joints (TJC and SJC)), demographics, blood profiles (hemoglobin, hematocrit, leukocyte- and thrombocytes levels) and ESR at baseline (Supplementary Table 6). Baseline was defined as the first visit to the clinic (set A & C) or the moment of inclusion in the trial (set B). Patients with missing lab or joint location variables were dropped (Supplementary Fig. 1).

We normalized the numerical data using a Yeo-Johnson transformation, except for the ESR levels where we applied a log transformation due to their log-normal distribution. For the categorical data, we implemented one-hot encoding, which created separate binary fields (yes/no) for each possible category value.

### Construction of patient embedding

We integrated the different EHR data types to create a condensed patient representation (called a patient embedding) using a multi-modal auto-encoder (MMAE). This MMAE had a narrowing structure with encoder layers containing 128, 64, and finally 8 neurons. For categorical data, we used sigmoid activation functions with Bernoulli loss, while numerical data utilized ReLU activation with Gaussian loss. To prevent overfitting, we compared performance between our training set (80% of data) and validation set (20%), using the Adam optimizer throughout.

After creating the patient embedding, we grouped patients into sub-communities based on clinical similarities using PhenoGraph<sup>56</sup>. We selected

PhenoGraph rather than more common methods like K-means clustering because it performs better with our sparse, high-dimensional medical data.

### Cluster interpretation

For each cluster, we examined the characteristics and visualized the phenotype on a pictorial mannequin with an integrated heatmap to demonstrate the number of involved joints. We used a surrogate ML-technique to model the cluster assignment and subjected this model to a SHAP (SHapley Additive exPlanations)<sup>57</sup> analysis to retrospectively identify the most important variables per cluster. The SHAP plots show the strength and direction of impact of that variable for each patient (also those who are not assigned to that cluster).

### Cluster validation

To confirm that our identified clusters comprised a stable and relevant partitioning, we performed a number of validation checks. We examined cluster stability by measuring how often patients co-cluster across 1000 random subsets of the data and assessed possible factors influencing the partitioning. Next, we conducted a Local Inverse Simpson's Index analysis<sup>58</sup> to assess whether physicians are evenly distributed across the clusters or if the patient subgroups merely reflect the reporting differences between physicians (see supplemental material).

To infer the clinical relevance, clinical outcomes were evaluated using a Cox regression model, including: time to MTX-failure (defined by replacement of- or adding an additional DMARD to MTX) and remission (DAS44 < 1.6) within 1 year. Moreover, we evaluated the replicability on an external dataset (set B & set C), where individuals are assigned to clusters in accordance with the previously learned patient embedding (see supplemental material).

### Diversity in synovial characteristics

In addition to the clinical validation, we compared the characteristics of the synovium across the four subgroups, by repeating the clustering analysis in set D (SYNGem cohort). Specifically, we compared the overall Krenn synovitis score (KSS)<sup>26</sup> and its individual components (lining layer hyperplasia, stromal density and inflammatory infiltrate). Each KSS component was measured on an ordinal scale (0: none, 1: low, 2: moderate, 3: severe). To analyze this, we used ordinal logistic regression, which estimates the odds of transitioning from one severity level to the next, thus relying on sufficient representation. However, since our study focused on treatment-naïve RA patients with active synovitis, the lowest level (0: none) for each KSS subitem was rarely observed. As a result, when present, it was combined with the 'low' category (1: low).

### Statistical tests

We used the Kruskal–Wallis test followed by Dunn's post-hoc test to compare numerical values across multiple groups. For survival analysis we used the log-rank test to examine the overall trend and a univariate Cox-regression<sup>59</sup> to quantify the cluster differences. We inferred the DAS-remission status during the survival analysis, carrying the last observation forward if it was missing (effectively the same as time to event). The proportional hazards assumption was verified by examining the Schoenfeld residuals<sup>60</sup>. We also adjusted the Cox regression model for MTX-response on covariates suspected to influence treatment response (e.g. ACPA positivity, RF positivity, SJC, TJC, Sex and Age). The statistical significance was inferred with ANOVA. For the ordinal values of the KSS components we used an ordered logit, followed up by a post-hoc Wald test to look at individual differences.

### Web Interface

All of our scripts are publicly available online at Github<sup>61</sup>. Moreover, we developed an interactive web tool (<https://knevel-lab.github.io/Rheumalyze/>) that enables users to map their cohort onto our patient embedding, allowing them to identify the JIP phenotypes present in their data.

### Data availability

The datasets generated and analyzed in this study are not publicly available due to strict privacy laws protecting patient data. However, the datasets may be made available by the corresponding author upon reasonable request.

### Code availability

We built an interactive webtool (<https://knevel-lab.github.io/Rheumalyze/>) that users can use to project their patients on the Leiden data, to infer the JIP-phenotype within their own dataset. We have made our scripts available in a public repository at GitHub<sup>61</sup>.

Received: 18 March 2025; Accepted: 8 September 2025;

Published online: 23 October 2025

### References

- Grassi, W. et al. The clinical features of rheumatoid arthritis. *Eur. J. Radiol.* **27**, S18–S24 (1998).
- Heidari, B. Rheumatoid arthritis: early diagnosis and treatment outcomes. *Casp. J. Intern. Med.* **2**, 161–170 (2011).
- Alivernini, S. et al. Inclusion of synovial tissue-derived characteristics in a nomogram for the prediction of treatment response in treatment-naïve rheumatoid arthritis patients. *Arthritis Rheumatol.* **73**, 1601–1613 (2021).
- Padyukov, L. Genetics of rheumatoid arthritis. *Semin. Immunopathol.* **44**, 47–62 (2022).
- Syversen, S. W. et al. Prediction of radiographic progression in rheumatoid arthritis and the role of antibodies against mutated citrullinated vimentin: results from a 10 year prospective study. *Ann. Rheum. Dis.* **69**, 345–351 (2010).
- Matthijssen, X. M. E. et al. Enhanced treatment strategies and distinct disease outcomes among autoantibody-positive and -negative rheumatoid arthritis patients over 25 years: a longitudinal cohort study in the Netherlands. *PLoS Med.* **17**, e1003296 (2020).
- Pedersen, M. et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res. Ther.* **8**, R133 (2006).
- Nordberg, L. B. et al. Patients with seronegative RA have more inflammatory activity compared with patients with seropositive RA in an inception cohort of DMARD-naïve patients classified according to the 2010 ACR/EULAR criteria. *Ann. Rheum. Dis.* **76**, 341–345 (2017).
- Sugihara, T. & Harigai, M. Targeting low disease activity in elderly-onset rheumatoid arthritis: current and future roles of biological disease-modifying antirheumatic drugs. *Drugs Aging* **33**, 97–107 (2016).
- Deal, C. L. et al. The clinical features of elderly-onset rheumatoid arthritis: a comparison with younger-onset disease of similar duration. *Arthritis Rheum.* **28**, 987–994 (1985).
- Ahlqvist, E. et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol.* **6**, 361–369 (2018).
- Howrylak, J. A. et al. Classification of childhood asthma phenotypes and long-term clinical responses to inhaled anti-inflammatory medications. *J. Allergy Clin. Immunol.* **133**, 1289–1300 (2014).
- Angelini, F. et al. Osteoarthritis endotype discovery via clustering of biochemical marker data. *Ann. Rheum. Dis.* **81**, 666–675 (2022).
- Lewis, M. J. et al. Molecular portraits of early rheumatoid arthritis identify clinical and treatment response phenotypes. *Cell Rep.* **28**, 2455–2470.e5 (2019).
- Liiso-Ribera, G. et al. Synovial tissue signatures enhance clinical classification and prognostic/treatment response algorithms in early inflammatory arthritis and predict requirement for subsequent biological therapy: results from the pathobiology of early arthritis cohort (PEAC). *Ann. Rheum. Dis.* **78**, 1642–1652 (2019).
- Jung, S. M., Park, K. S. & Kim, K. J. Clinical phenotype with high risk for initiation of biologic therapy in rheumatoid arthritis: a data-driven cluster analysis. *Clin. Exp. Rheumatol.* **39**, 1282–1290 (2021).

17. Curtis, J. R. et al. Data-driven patient clustering and differential clinical outcomes in the brigham and women's rheumatoid arthritis sequential study registry. *Arthritis Care Res. (Hoboken)* **73**, 471–480 (2021).
18. Merola, J. F., Espinoza, L. R. & Fleischmann, R. Distinguishing rheumatoid arthritis from psoriatic arthritis. *RMD Open* **4**, e000656 (2018).
19. Onken, A., Yague, J. G. & Sakata, S. Deep multimodal autoencoders for identifying latent representations of spike counts and local field potentials. *Bernstein Conf.* <https://doi.org/10.12751/nncn.bc2018.0223> (2018).
20. Steinbach, M., Ertöz, L. & Kumar, V. The challenges of clustering high-dimensional data. In *New Directions in Statistical Physics: Econophysics, Bioinformatics, and Pattern Recognition* (ed. Wille, L. T.) 273–309 (Springer, Berlin, 2004).
21. Cai, J. et al. Efficient deep embedded subspace clustering. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* 1–10 (IEEE, 2022).
22. Castela Forte, J. et al. Identifying and characterizing high-risk clusters in a heterogeneous ICU population with deep embedded clustering. *Sci. Rep.* **11**, 12109 (2021).
23. Nishimura, K. et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Ann. Intern. Med.* **146**, 797–808 (2007).
24. van Nies, J. A. B. et al. Evaluating relationships between symptom duration and persistence of rheumatoid arthritis: does a window of opportunity exist? results on the Leiden early arthritis clinic and ESPOIR cohorts. *Ann. Rheum. Dis.* **74**, 806–812 (2015).
25. Visser, H. et al. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum.* **46**, 357–365 (2002).
26. Krenn, V. et al. Synovitis score: discrimination between chronic low-grade and high-grade synovitis. *Histopathology* **49**, 358–364 (2006).
27. Ciurea, A., Ospelt, C. Location-specific treatment of chronic inflammatory joint disease. *Nat. Rev. Rheumatol.* **21**, 507–508 (2025).
28. Lee, S. W. et al. Prevalence of feet and ankle arthritis and their impact on clinical indices in patients with rheumatoid arthritis: a cross-sectional study. *BMC Musculoskelet. Disord.* **20**, 420 (2019).
29. Hoque, A., Steultjens, M., Dickson, D. M. & Hendry, G. J. Patients' and clinicians' perspectives on the clinical utility of the rheumatoid arthritis foot disease activity index. *Rheumatol. Int.* **42**, 1807–1817 (2022).
30. Wechalekar, M. D. et al. Active foot synovitis in patients with rheumatoid arthritis: applying clinical criteria for disease activity and remission may result in underestimation of foot joint involvement. *Arthritis Rheum.* **64**, 1316–1322 (2012).
31. Ciurea, A. et al. Joint-level responses to tofacitinib and methotrexate: a post hoc analysis of data from ORAL Start. *Arthritis Res. Ther.* **25**, 185 (2023).
32. Ai, R. et al. Joint-specific DNA methylation and transcriptome signatures in rheumatoid arthritis identify distinct pathogenic processes. *Nat. Commun.* **7**, 11849 (2016).
33. Buckley, C. D. et al. Location, location, location: how the tissue microenvironment affects inflammation in RA. *Nat. Rev. Rheumatol.* **17**, 195–212 (2021).
34. Frank-Bertoncelj, M. et al. Epigenetically-driven anatomical diversity of synovial fibroblasts guides joint-specific fibroblast functions. *Nat. Commun.* **8**, 14852 (2017).
35. van der Helm-van Mil, A. H. M. et al. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. *Arthritis Res. Ther.* **7**, R949–R958 (2005).
36. Bugatti, S. et al. The clinical value of autoantibodies in rheumatoid arthritis. *Front. Med. (Lausanne)* **5**, 339 (2018).
37. Aletaha, D. et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum.* **62**, 2569–2581 (2010).
38. Burgers, L. E. et al. Differences in the symptomatic phase preceding ACPA-positive and ACPA-negative RA: a longitudinal study in arthralgia during progression to clinical arthritis. *Ann. Rheum. Dis.* **76**, 1751–1754 (2017).
39. Sokolova, M. V., Schett, G. & Steffen, U. Autoantibodies in rheumatoid arthritis: historical background and novel findings. *Clin. Rev. Allergy Immunol.* **63**, 138–151 (2022).
40. Duong, S. Q. et al. Clinical predictors of response to methotrexate in patients with rheumatoid arthritis: a machine learning approach using clinical trial data. *Arthritis Res. Ther.* **24**, 162 (2022).
41. Nordberg, L. B. et al. Comparing the disease course of patients with seronegative and seropositive rheumatoid arthritis fulfilling the 2010 ACR/EULAR classification criteria in a treat-to-target setting: 2 year data from the ARCTIC trial. *RMD Open* **4**, e000752 (2018).
42. Naides, S. J. et al. Rheumatologic manifestations of human parvovirus B19 infection in adults: initial two-year clinical experience. *Arthritis Rheum.* **33**, 1297–1309 (1990).
43. Takahashi, Y. et al. Human parvovirus B19 as a causative agent for rheumatoid arthritis. *Proc. Natl. Acad. Sci. USA* **95**, 8227–8232 (1998).
44. Burns, T. M. & Calin, A. The hand radiograph as a diagnostic discriminant between seropositive and seronegative 'rheumatoid arthritis': a controlled study. *Ann. Rheum. Dis.* **42**, 605–612 (1983).
45. Rivelles, F. et al. Rituximab versus tocilizumab in rheumatoid arthritis: synovial biopsy-based biomarker analysis of the phase 4 R4RA randomized trial. *Nat. Med.* **28**, 1256–1268 (2022).
46. Heckert, S. L. et al. Joint inflammation tends to recur in the same joints during the rheumatoid arthritis disease course. *Ann. Rheum. Dis.* **81**, 169–174 (2022).
47. Baker, J. F. et al. Associations between body mass, radiographic joint damage, adipokines and risk factors for bone loss in rheumatoid arthritis. *Rheumatology* **50**, 2100–2107 (2011).
48. George, M. D. et al. Impact of obesity and adiposity on inflammatory markers in patients with rheumatoid arthritis. *Arthritis Care Res.* **69**, 1789–1798 (2017).
49. Mohamed-Ali, V., Pinkney, J. H. & Coppack, S. W. Adipose tissue as an endocrine and paracrine organ. *Int. J. Obes. Relat. Metab. Disord.* **22**, 1145–1158 (1998).
50. Visser, M., Bouter, L. M., McQuillan, G. M., Wener, M. H. & Harris, T. B. Elevated C-reactive protein levels in overweight and obese adults. *JAMA* **282**, 2131–2135 (1999).
51. Bauer, E. M. et al. Joint-specific assessment of swelling and power Doppler in obese rheumatoid arthritis patients. *BMC Musculoskelet. Disord.* **18**, 99 (2017).
52. Zhang, F. et al. Deconstruction of rheumatoid arthritis synovium defines inflammatory subtypes. *Nature* **623**, 616–624 (2023).
53. Maarseveen, T. D. et al. Machine learning electronic health record identification of patients with rheumatoid arthritis: algorithm pipeline development and validation study. *JMIR Med. Inform.* **8**, e23930 (2020).
54. Maarseveen, T. D. et al. Handwork vs machine: a comparison of rheumatoid arthritis patient populations as identified from EHR free-text by diagnosis extraction through machine-learning or traditional criteria-based chart review. *Arthritis Res. Ther.* **23**, 174 (2021).
55. Heimans, L. et al. A two-step treatment strategy trial in patients with early arthritis aimed at achieving remission: the IMPROVED study. *Ann. Rheum. Dis.* **73**, 1356–1361 (2014).
56. Levine, J. H. et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).
57. Roth, A. E. *The Shapley Value: Essays in Honor of Lloyd S. Shapley* (Cambridge Univ. Press, 1988).
58. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat. Methods* **16**, 1289–1296 (2019).
59. Cox, D. R. Regression models and life-tables. *J. R. Stat. Soc. Ser. B* **34**, 187–202 (1972).
60. Schoenfeld, D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika* **67**, 145–153 (1980).

61. Maarseveen, T. D. *EHR Clustering RA*. <https://github.com/levrex/EHR-Clustering-RA> (2022).

## Acknowledgements

This project has received funding from two large European Union's Horizon grants for Europe research and innovation. First for SQUEEZE (activity No. 101095052) and secondly for SPIDeRR (activity No. 101080711). There was additional financial support from the European research council for the GlycanSwitch project (activity No. 101071386). This study was also co-founded by the ZonMW Klinische Fellow No. 40-00703-97-19069, as well as the Zonmw Open Competitie, No. 09120012110075. We would like to express our gratitude to Nick Bos for developing a web tool to cluster patients. Our thanks also go to Samantha Jurado-Zapata and David Steeman for their assistance with extracting and processing Electronic Health Record data from the Leiden University Medical Center. Furthermore, we deeply appreciate Bas van der Wal for his efforts in preparing the Reumazorg Zuidwest Nederland data and Joy van der Pol for his work on preparing the IMPROVED trial data.

## Author contributions

R.K. and T.M. developed the study design together with E.B.v.d.A. and M.P.M. T.M. ran the cluster analysis, while S.B. ran the permutation analysis. B.B.d.K. repeated the survival analysis in the replication set. S.A.B. provided the replication data of the IMPROVED trial, while K.G. and J.V.v.D. provided the data from Reumazorg Zuid West Nederland. S.A., L.A.C. and S.P. provided and processed the SYNGem cohort data for the histological analysis. All authors contributed to the interpretation of the results, and provided ideas for further downstream analysis. A.H.M. annotated the classification criteria. R.K. and T.M. drafted the first version of the manuscript. All authors reviewed and approved the final draft of the manuscript to be submitted.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-01997-1>.

**Correspondence** and requests for materials should be addressed to Rachel Knevel.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025