

<https://helda.helsinki.fi>

---

## Cluster analysis identifies unmet healthcare needs among patients with rheumatoid arthritis

Mars, Nina

2022

---

Mars , N , Kerola , A M , Kauppi , M J , Pirinen , M , Elonheimo , O & Sokka-Isler , T 2022 , ' Cluster analysis identifies unmet healthcare needs among patients with rheumatoid arthritis ' p̈y , Scandinavian Journal of Rheumatology , vol. 51 , no. 5 , pp. 355 362

---

<http://hdl.handle.net/10138/349873>

<https://doi.org/10.1080/03009742.2021.1944306>

---

cc\_by

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*



## Cluster analysis identifies unmet healthcare needs among patients with rheumatoid arthritis

N Mars, AM Kerola, MJ Kauppi, M Pirinen, O Elonheimo & T Sokka-Isler

To cite this article: N Mars, AM Kerola, MJ Kauppi, M Pirinen, O Elonheimo & T Sokka-Isler (2022) Cluster analysis identifies unmet healthcare needs among patients with rheumatoid arthritis, Scandinavian Journal of Rheumatology, 51:5, 355-362, DOI: [10.1080/03009742.2021.1944306](https://doi.org/10.1080/03009742.2021.1944306)

To link to this article: <https://doi.org/10.1080/03009742.2021.1944306>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 13 Sep 2021.



[Submit your article to this journal](#)



Article views: 917



[View related articles](#)



[View Crossmark data](#)

# Cluster analysis identifies unmet healthcare needs among patients with rheumatoid arthritis

N Mars<sup>1,2</sup>, AM Kerola<sup>1,3</sup>, MJ Kauppi<sup>3,4</sup>, M Pirinen<sup>2,5,6</sup>, O Elonheimo<sup>1</sup>, T Sokka-Isler<sup>7</sup>

<sup>1</sup>Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>2</sup>Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland

<sup>3</sup>Department of Internal Medicine, Päijät-Häme Joint Authority for Health and Wellbeing, Lahti, Finland

<sup>4</sup>Clinicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>5</sup>Department of Public Health, University of Helsinki, Helsinki, Finland

<sup>6</sup>Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

<sup>7</sup>Department of Rheumatology, Jyväskylä Central Hospital, Jyväskylä, Finland

**Objective:** To identify the patterns of healthcare resource utilization and unmet needs of persistent disease activity, pain, and physical disability in rheumatoid arthritis (RA) by cluster analysis.

**Method:** Patients attending the Jyväskylä Central Hospital rheumatology unit, Finland, were, from 2007, prospectively enrolled in a clinical database. We identified all RA patients in 2010–2014 and combined their individual-level data with well-recorded administrative data on all public healthcare contacts in fiscal year 2014. We ran agglomerative hierarchical clustering (Ward's method), with 28-joint Disease Activity Score with three variables, Health Assessment Questionnaire index, pain (visual analogue scale 0–100), and total annual health service-related direct costs (€) as clustering variables.

**Results:** Complete-case analysis of 939 patients derived four clusters. Cluster C1 (remission and low costs, 550 patients) comprised relatively young patients with low costs, low disease activity, and minimal disability. C2 (chronic pain, disability, and fatigue, 269 patients) included those with the highest pain and fatigue levels, and disability was fairly common. C3 (inflammation, 97 patients) had rather high mean costs and the highest average disease activity, but lower average levels of pain and less disability than C2, highlighting the impact of effective treatment. C4 (comorbidities and high costs, 23 patients) was characterized by exceptionally high costs incurred by comorbidities.

**Conclusions:** The majority of RA patients had favourable outcomes and low costs. However, a large group of patients was distinguished by chronic pain, disability, and fatigue not unambiguously linked to disease activity. The highest healthcare costs were linked to high disease activity or comorbidities.

For patients with rheumatoid arthritis (RA), the advent of more effective anti-rheumatic treatments and the principle of aiming at remission or low disease activity has translated into improved outcomes, such as better control of disease activity, functional capacity, and lower healthcare resource utilization (1–3). Despite the therapeutic advances, for many patients with RA, poor physical functioning, pain, and comorbidity remain persistent problems (4). This has led to the recognition of patient-reported outcomes such as pain, fatigue, and functional status as unmet needs in RA care (5). To further improve the care of RA patients and to facilitate the optimal allocation of limited resources, we need to

better identify the patterns by which these symptoms occur and to quantify the proportion of RA patients affected.

RA is associated with a multitude of comorbidities, which, in turn, are associated with poor health-related quality of life and higher healthcare expenditures (6). Many prior studies on economic consequences of rheumatoid arthritis (RA) have focused particularly on overall costs, cost components, and medication costs (7–9). Previously, we reported that 10% of RA patients use as many healthcare resources as the remaining 90%, and that these high utilizers were characterized by costly comorbidities and higher levels of chronic pain (10). Although healthcare costs in RA may be influenced by disease activity and functional status (11, 12), few data exist on patterns of patient-reported outcomes and healthcare costs for patients with RA.

Cluster analysis is an exploratory statistical method used in many fields for grouping based on similarity (13), but it is a novel approach in research on

N Mars, Faculty of Medicine and Institute for Molecular Medicine Finland, University of Helsinki, P.O. Box 20 (Tukholmankatu 8), FI-00014, Finland.

E-mail: [nina.mars@helsinki.fi](mailto:nina.mars@helsinki.fi)

Accepted 14 June 2021

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

DOI: 10.1080/03009742.2021.1944306

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

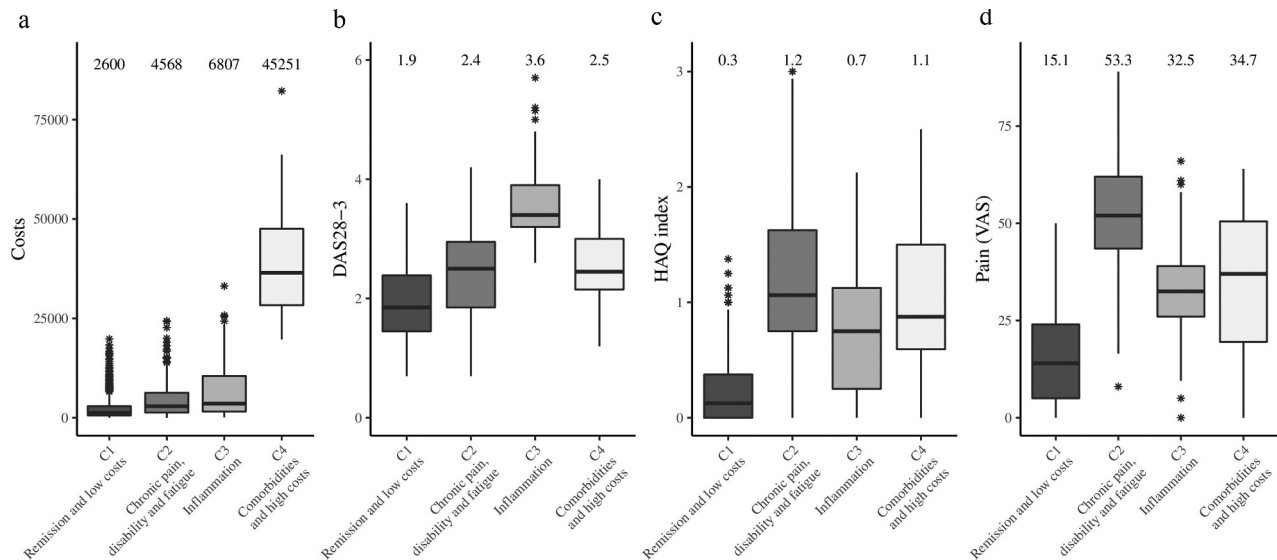


Figure 1. Boxplots representing the distributions of the clustering variables for each cluster. The variables were (A) costs, (B) disease activity measured by the Disease Activity Score in 28 joints with three variables (DAS28-3), (C) disability measured by the Health Assessment Questionnaire (HAQ) index, and (D) pain measured by the visual analogue scale (VAS). Means are shown above the boxplots. For individual patients, the median of time-dependent clinical variables was considered. The black line is the median, the box represents the interquartile range (IQR = Q3 – Q1), the lower whisker is Q1 – 1.5 \* IQR, and the upper whisker is Q3 + 1.5 \* IQR. Stars represent outlier values located outside the whiskers.

rheumatic diseases. Using cluster analysis, we set out to explore the patterns of healthcare resource utilization and unmet needs of pain, disability, and persistent disease activity in RA. As cluster analysis requires high-quality data as input, this study was carried out by linking two high-quality registers: a population-based data set from a rheumatology clinic, which involves repeated measures on both disease activity and patient-reported outcomes, and data on healthcare resource utilization.

## Method

We identified adult (age  $\geq 18$  years) patients who lived in the Jyväskylä area (population approximately 140 000) and who had visited the rheumatology clinic at Finland's largest non-university hospital [Jyväskylä Central Hospital (JCH)] between 2010 and 2014. A structured digital database (GoTreatIT® Rheuma application, DiaGraphIT) systematically collected individual-level clinical data as part of the electronic medical records (14) between May 2007 and 16 March 2016. Each patient completes a comprehensive questionnaire including measures of pain, fatigue, and functional status before every visit to JCH rheumatology clinic.

The healthcare utilization data involved a system similar to diagnosis-related group (DRG), one suitable for both inpatient and outpatient care. This was used for grouping all the RA patients' diagnoses for fiscal year 2014, and for estimating the respective health service-related direct costs (€; price level for 2014). The cost estimation tool acknowledges disease category, age, gender, healthcare unit and provider, and procedures,

and comprises all public healthcare contacts: both primary and speciality care, inpatient and outpatient care, the emergency department, and contacts with all healthcare professionals (physicians, nurses, and rehabilitation workers). Additional details of both data sets have been described previously (10). We combined the data sets using the unique Finnish national identification numbers, selecting RA patients with at least one healthcare contact in 2014. As healthcare utilization data were obtained for 2014, our inclusion criteria were patients diagnosed with RA before or in 2014, who had visits to the rheumatology clinic within 5 years prior (2010–2014) to collection of cost data. To capture patterns of persistent disease activity, pain, and physical disability, we used all individual-level clinical data available for these patients within the registry (2007–2016).

The clustering variables were Disease Activity Score with three variables based on 28-joint count–erythrocyte sedimentation rate (DAS28-3-ESR) for disease activity, the Health Assessment Questionnaire disability index (HAQ, 0–3) for disability, the visual analogue scale for pain (VAS, 0–100, during the past week), and total annual health service-related direct costs (€; referred to as 'costs'). For individual patients, we considered the median of time-dependent clinical variables. Correlations for evaluating alternative summary metrics [mean and area under the curve (AUC) of individual trajectories calculated using the trapezoid rule] are shown in [Supplementary Figure 1](#). We also explored the effect of replacing the variable on costs with the Rheumatic Disease Comorbidity Index (RDCI) (15) with comorbidities reported in the clinical data set.

Our main clustering algorithm was agglomerative hierarchical clustering (AHC), with secondary analyses

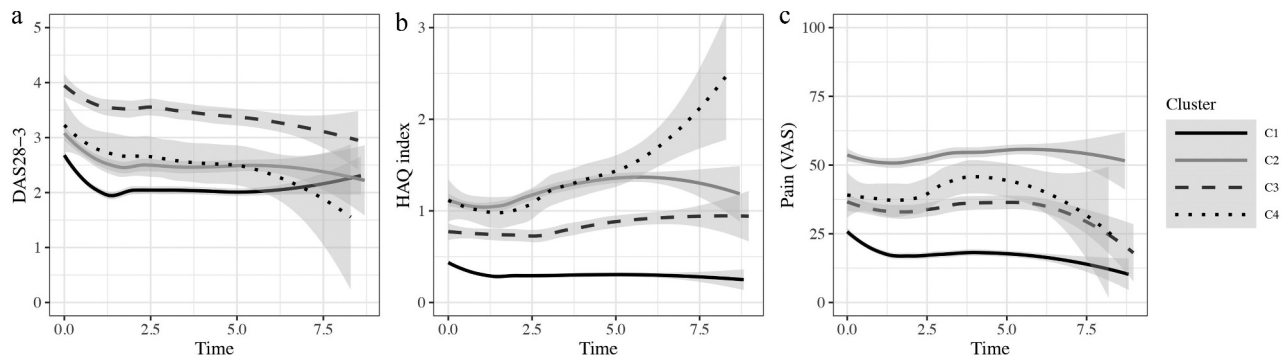


Figure 2. Longitudinal patterns by cluster for (A) disease activity measured by Disease Activity Score in 28 joints with three variables (DAS28-3), (B) disability measured by the Health Assessment Questionnaire (HAQ) index, and (C) pain measured by the visual analogue scale (VAS). The x-axis represents years of follow-up after the diagnosis. Curves are locally estimated scatterplot smoothing (LOESS) curves, fitted by cluster to individual trajectories over the follow-up [R package ggplot, function `stat_smooth()` with method parameter 'loess']. The grey areas represent 95% confidence intervals.

performed with k-means clustering. AHC starts with each individual in their own group and proceeds step by step by merging the two groups closest to each other according to the similarity measure. After taking the square root of costs (and RDCI for analyses applying RDCI) and scaling all the variables to zero mean and unit variance, we ran AHC by Ward's method (Ward2 algorithm), with the similarity measure defined by the Euclidean distance, following the approach of some earlier publications (16, 17). As a sensitivity analysis of cluster stability, we ran k-means clustering (R package `fpc`, function `kmeansruns` with 100 runs).

The number of clusters chosen was based on examination of the dendrogram, and by assessing the cluster number by three metrics: the average silhouette width, the within-cluster sum of squares by cluster, and the gap statistic (Supplementary Figure 2). With all metrics, the optimal number of clusters was two, but this would have distinguished only patients who are doing well with respect to our clinical characteristics, and those who are not. Based on the dendrogram, and supported by the three metrics, we therefore chose four clusters to explore the patterns of unmet needs. Cluster labels were assigned by examining the distributions of the cluster variables. To visualize the overlap between clusters, we performed the standard principal component analysis (PCA). PCA creates new, uncorrelated variables (principal components) as linear combinations of the original variables. Typically, a few leading principal components will reveal the main structure of the data.

We describe the differences among the clusters for patient-reported outcome measures, health service-related costs, and comorbidities systematically recorded in the clinical data. Medication data are reported as ever- and never-users for disease-modifying anti-rheumatic drugs (DMARDs), biological disease-modifying anti-rheumatic drugs (bDMARDs), and methotrexate. Fatigue was rated on the VAS (0–100)

during the previous week of answering the questionnaire before each visit. For comparing categorical variables, we used Fisher's exact test, and for continuous variables, the Kruskal–Wallis test.

No patients were involved in planning or setting the research questions, or when interpreting the results. Analyses were performed using R version 3.4.0. In Finland, linkage of registry-based data requires no ethics approval or patient consent. The study was approved by the local medical records administrator (JCH).

## Results

There were 939 individuals meeting the selection criteria with no missing data with respect to the clustering variables. A comparison of individuals included and excluded is given in Supplementary Table 1. The mean number of visits was 8.9 [median 7.0, interquartile range (IQR) 4.0–11.0]. The mean duration of follow-up was 2.6 years (median 2.1 years, IQR 0.7–4.2 years, maximum 9.0 years). We considered four clusters, C1–C4 (Table 1; dendrogram in Supplementary Figure 3) ordered by increasing total healthcare costs. Distributions for DAS28-3, HAQ index, and pain are shown in Figure 1. The longitudinal patterns for DAS28-3, HAQ index, and pain by cluster are shown in Figure 2. Labelled 'remission and low costs', C1 was the largest cluster, with 550 patients. It constituted relatively young patients with low costs, low disease activity, minimal disability, and the lowest number of comorbidities. Labelled 'chronic pain, disability, and fatigue', C2, with 269 patients, was characterized by the highest pain levels (VAS group mean 53.3), and disability was fairly common. In all groups, fatigue was common, but C2 ('chronic pain, disability, and fatigue') had the highest group mean of 50.6 (Table 1).

Table 1. Patient characteristics.

	C1	C2	C3	C4	p*
N (%)	550 (58.6)	269 (28.6)	97 (10.3)	23 (2.4)	
Age (years), mean ± sd	58.7 ± 14.9	66.6 ± 12.9	66.8 ± 13.7	69.2 ± 13.5	< 0.0001
Women (%)	70.4	72.9	69.1	73.9	0.84
Disease duration (years), mean ± sd	12.2 ± 8.6	16.7 ± 12.5	15.8 ± 11.9	15.8 ± 11.3	0.0001
Patients with early RA (%)†	9.8	8.3	8.5	8.7	0.91
Seropositive (%)	71.5	68.8	76.3	82.6	0.57
Erosions (%)	45.6	46.8	53.6	52.2	0.24
Fatigue, mean ± sd	20.0 ± 18.7	50.6 ± 20.6	32.1 ± 19.2	33.1 ± 23.7	< 0.0001
Ever DMARDs (%)	97.6	95.9	97.9	100.0	1.00
Ever methotrexate (%)	91.1	89.6	93.8	100.0	0.63
Ever bDMARDs (%)	22.7	24.5	46.4	39.1	< 0.0001
Mean number of comorbidities in clinical data	2.1	3.7	3.5	5.2	< 0.0001
Mean number of visits to rheumatology unit†	8.5	8.8	11.4	11.0	0.06
Mean total costs/patient (€)	2600	4568	6807	45 251	
Mean rheumatic disease costs/patient (€)	1198	1583	2694	4075	< 0.0001
Mean speciality care inpatient costs/patient (€)	518	937	1853	22 740	< 0.0001
Total costs in 2014 (million €)	1.43	1.22	0.66	1.04	

For continuous variables, the means of individual patients' medians are presented. Clusters C1–C4 are ordered based on total costs. RA, rheumatoid arthritis; DMARD, disease-modifying anti-rheumatic drug; bDMARD, biological disease-modifying anti-rheumatic drug. \*The p-value represents a global comparison across the four clusters. †Diagnosis in 2013 or 2014. ‡As of 2007.

The most characteristic feature in C3 ('inflammation', 97 patients) was rather high mean costs. C3 also had the highest average DAS28-3. Despite high disease activity compared to that of C2, C3 showed lower average levels of pain and less disability. All clusters showed high variability in DAS28-3, suggesting that all clusters comprised patients with varying levels of disease activity.

With only 23 patients, C4 ('comorbidities and high costs') was small, heterogeneous, and characterized by exceptionally high average costs. The main cost driver in C4 was costly and severe comorbidities. Detailed evaluation revealed that these comprised mainly malignancies, severe gastrointestinal diseases such as gastrointestinal bleeding, severe infections, and complications of joint replacement surgery.

In the clinical data, 769 patients (81.9%) had at least one comorbidity. Hypertension, hypercholesterolaemia, osteoporosis, and hypothyroidism were common in all clusters (Supplementary Table 2). In C2, 13.4% had physician-diagnosed fibromyalgia. For C3 (inflammation), coronary artery disease was present in 10.3%, and for C2 (chronic pain, disability, and fatigue) in 8.6%. The top comorbidities by annual health service-related costs are given in Supplementary Table 3.

Next, we performed several sensitivity analyses. Clustering with RDCI replacing costs and without either costs or RDCI yielded overall similar results for clusters, particularly with respect to C1 and C2, but not all patients were assigned to a similar cluster by the different clustering variables, and cluster sizes were more uniform in size (Supplementary Figures 4 and 5). Clustering with RDCI failed to identify a clear pattern for inflammation for C3 (Supplementary Figure 4). Clustering without either costs or RDCI yielded highly similar patterns for C1–C3 as our main clustering results, but was unable to distinguish individuals with respect to costs (Supplementary Figure 5). k-Means clustering produced overall highly similar clusters to hierarchical cluster analysis, and the same labels were assignable for C1–C4 (Supplementary Figure 6, Supplementary Table 4). The cluster sizes remained consistent, with 401 individuals (42.7%) in C1, 206 (21.9%) in C2, 257 (27.4%) in C3, and 75 (8.0%) in C4, with the most substantial change being that 29.1% of individuals in hierarchical clustering cluster C1 (remission and low costs) moved with k-means to the cluster representing inflammation. Moreover, 195 of the 939 individuals (20.8%) were assigned to cluster C2 (chronic pain, disability, and fatigue) with both algorithms.

Lastly, to visualize the overlap between clusters, we performed standard PCA. PCA demonstrated that C3 (inflammation) shared similarities with all the other clusters, and overall, cluster overlap was apparent (Supplementary Figure 7). The most widespread was C4 (comorbidities and high costs), demonstrating the heterogeneity of patients in C4.

## Discussion

Effective treatment of RA seems to promise positive outcomes for the majority: the largest cluster, C1, comprising over half of the patients, had favourable outcomes and low average costs. However, the second largest cluster, comprising approximately one-third of patients, showed unmet needs of pain, disability, and fatigue (cluster C2). These patients were characterized by substantial self-reported pain and fatigue, and disability not explicitly linked to disease activity. The pain and fatigue measurements consisted of individual medians, suggesting that many suffer from chronic pain, and fibromyalgia was diagnosed in 13.4%.

One-tenth of all patients (cluster C3) had characteristics suggesting chronic inflammation, and some of these patients also displayed moderate to high levels of pain and disability. Despite the inflammation, many seem to have maintained good physical functioning, perhaps as a result of active treatment. Of note, 46% of patients in C3 had used bDMARDs at some point in their disease course. A clinically important observation is that in this cluster with persistent disease activity, some of the patients reported low levels of pain, and the highest levels of pain were found not in this cluster, but in C2. Pain that is not correlated with levels of disease activity is an important challenge in treating patients with rheumatic diseases, and our results indicate that this is an unmet need in nearly every third patient with RA. Despite similar age, disease duration, and number of comorbidities, patients in C3 had more visits to the rheumatology unit and higher total and rheumatic disease-related costs than C2, implying that more resources are allocated to the care of those with high disease activity than of those with chronic pain and disability.

Over the past few decades, the introduction of more aggressive treatment strategies for RA has translated into reductions in disease activity and inflammatory markers, while patient-reported outcomes such as pain, fatigue, functional disability, and mental health remain to be similarly improved (18, 19). Recent reviews and guidelines recognize pain and physical functioning as key domains of unmet needs in RA and emphasize the importance of differentiating localized and generalized pain (4, 20). Fatigue is generally associated with a high comorbidity burden, disease activity, and disability (21), but may also be linked to chronic pain. Overall, improving the management of chronic pain is key, and these patients may also benefit from reinforced support by a multiprofessional team.

The largest cluster, C1, comprising over half of the patients, had both low total costs and low rheumatic disease-related costs, and were doing well overall. They had the lowest number of comorbidities, low disease activity, and favourable patient-reported outcomes. The smallest cluster, C4, comprised patients with severe

comorbidities incurring high costs. Their comorbidity spectrum, fairly high average level of DAS28-3, and 40% having used bDMARDs all imply that the index disease may have contributed to their comorbidity burden and healthcare resource utilization. This finding is in line with our previous study showing that costly comorbidities account for most healthcare costs among those RA patients who utilize the most healthcare resources (10).

Prior studies from the USA have established clusters with characteristics similar to ours. One study identified a large cluster comprising the least ill patients and a small cluster with a high comorbidity burden (2). Another established clusters similar to our C1–C3, and identified patients with high levels of pain who displayed minimal signs of inflammation and manifested symptoms indicative of a chronic widespread pain syndrome (22). One study, using latent class analysis to identify comorbidity clusters among patients with rheumatic diseases, identified as the most prevalent patterns for RA cardiovascular disease and related risk factors, osteoporotic fractures, and depression (23). All of these, except for depression, were also highly prevalent across our clusters.

Recognizing pain, fatigue, and disability as common and persistent problems among RA patients, our study replicates previous findings using a population-based RA cohort in a European country with universal public healthcare. This study shares the limitations of other administrative-data studies, such as coding errors, but direct linkage to rheumatologist-validated clinical data and categorization of comorbidities reduce these biases. We deliberately examined only health service-related direct costs. The treatment strategies in JCH follow the Finnish current care guideline for RA (24), and we believe that the generalizability of our results on a national level is good. In terms of disease activity and severity of RA, the QUEST-RA study reports significant variation between countries, with lower-than-average disease activity, joint counts, pain, and HAQ in Finland (25), which may limit the generalizability of our results outside Finland.

In general, cluster analysis provides additional information beyond that captured by traditional research frameworks. Like many clusters in other medical specialities (26), our clusters make sense from a clinical perspective. Our clusters did, however, display heterogeneity. This heterogeneity is likely to arise from the relevant clinical characteristics being highly correlated. Heterogeneity may be reduced by increasing the number of clusters, which would, however, make it harder to identify and quantify key patterns. The clearest cluster division would have been between those with favourable outcomes (C1) and the rest with unmet needs (C2, C3, and C4 combined), but these two main clusters do not yet identify the specific domains of unmet needs.

Although different clustering–variable combinations and clustering algorithms showed some reclassification, the unmet needs remained evident with the different approaches. The most important cluster with respect to unmet needs, cluster C2 (chronic pain, disability, and fatigue), showed the highest stability between the different algorithms, and clustering could be used for identifying such patients in rheumatology clinics. Although we believe that our results are clinically relevant and useful in emphasizing the extent of pain, fatigue, and disability as unmet needs, the operationalization of our clusters would require replication in other RA cohorts. Moreover, data on costs are rarely available, and, in particular, our results on clustering with RDCI and without either RDCI or costs show that patient characteristics (at least explored by clustering) alone are insufficient to stratify individuals according to costs.

## Conclusion

We identified four clusters in RA based on patient-reported outcomes and healthcare resource utilization. Most patients were doing well with their disease and had low costs, in part reflecting the effects of modern anti-rheumatic treatment, but around one-third displayed an unmet need by presenting with chronic pain and fatigue not explicitly linked to disease activity. In clinical practice, it is important to acknowledge these clusters, which may also help to optimize treatment strategies. Alongside implementation of early and active RA treatment strategies, we highlight the importance of improving multiprofessional care of patients with chronic pain and disability.

## Acknowledgements

We thank Santeri Huvinen and Tuuli Pajunen for support in accessing the administrative data, and Carolyn Brimley Norris for language editing.

## Data availability

The data underlying this study cannot be shared owing to patient privacy.

## Disclosure statement

AMK has received speaker fees from Boehringer-Ingelheim, has attended advisory boards of Pfizer, Gilead, and Boehringer-Ingelheim, and received congress sponsorship from Pfizer, Celgene, UCB Pharma, Mylan, and Roche, outside the submitted work. The other authors report no potential conflicts of interest.



## Funding

This work was supported by a research grant from Finnish Research Foundation for Rheumatic Diseases (to NM) and by the Academy of Finland (grant number 331671 to NM). AMK has been supported by a research grant from FOREUM Foundation for Research in Rheumatology.

## References

- Alemao E, Joo S, Kawabata H, Al MJ, Allison PD, Rutten-van Molken MP, et al. Effects of achieving target measures in rheumatoid arthritis on functional status, quality of life, and resource utilization: analysis of clinical practice data. *Arthritis Care Res (Hoboken)* 2016;68:308–17.
- Curtis JR, Chen L, Greenberg JD, Harrold L, Kilgore ML, Kremer JM, et al. The clinical status and economic savings associated with remission among patients with rheumatoid arthritis: leveraging linked registry and claims data for synergistic insights. *Pharmacoepidemiol Drug Saf* 2017;26:310–9.
- Rantalaiho V, Puolakka K, Korpela M, Hannonen P, Mottonen T. Long-term results of the FIN-RACo trial; treatment with a combination of traditional disease-modifying anti-rheumatic drugs is an excellent option in early rheumatoid arthritis. *Clin Exp Rheumatol* 2012;30:S27–31.
- Taylor PC, Moore A, Vasilescu R, Alvir J, Tarallo M. A structured literature review of the burden of illness and unmet needs in patients with rheumatoid arthritis: a current perspective. *Rheumatol Int* 2016;36:685–95.
- Fautrel B, Alten R, Kirkham B, de la Torre I, Durand F, Barry J, et al. Call for action: how to improve use of patient-reported outcomes to guide clinical decision making in rheumatoid arthritis. *Rheumatol Int* 2018;38:935–47.
- An J, Nyarko E, Hamad MA. Prevalence of comorbidities and their associations with health-related quality of life and healthcare expenditures in patients with rheumatoid arthritis. *Clin Rheumatol* 2019;38:2717–26.
- Husberg M, Davidson T, Hallert E. Non-medical costs during the first year after diagnosis in two cohorts of patients with early rheumatoid arthritis, enrolled 10 years apart. *Clin Rheumatol* 2017;36:499–506.
- Huscher D, Mittendorf T, von Hinuber U, Kotter I, Hoese G, Pfafflin A, et al. Evolution of cost structures in rheumatoid arthritis over the past decade. *Ann Rheum Dis* 2015;74:738–45.
- van den Akker-van Marle ME, Chorus AM, Vliet Vlieland TP, van den Hout WB. Cost of rheumatic disorders in the Netherlands. *Best Pract Res Clin Rheumatol* 2012;26:721–31.
- Mars NJ, Kerola AM, Kauppi MJ, Pirinen M, Elonheimo O, Sokka-Isler T. Patients with rheumatic diseases share similar patterns of healthcare resource utilization. *Scand J Rheumatol* 2019;48:300–7.
- Huscher D, Merkesdal S, Thiele K, Zeidler H, Schneider M, Zink A, et al. Cost of illness in rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis and systemic lupus erythematosus in Germany. *Ann Rheum Dis* 2006;65:1175–83.
- Beresniak A, Gossec L, Goupille P, Saraux A, Bamberger M, Bregman B, et al. Direct cost-modeling of rheumatoid arthritis according to disease activity categories in France. *J Rheumatol* 2011;38:439–45.
- Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett* 2010;31:651–66.
- Vare P, Nikiphorou E, Hannonen P, Sokka T. Delivering a one-stop, integrated, and patient-centered service for patients with rheumatic diseases. *SAGE Open Med* 2016;4:2050312116654404.
- England BR, Sayles H, Mikuls TR, Johnson DS, Michaud K. Validation of the rheumatic disease comorbidity index. *Arthritis Care Res (Hoboken)* 2015;67:865–72.
- Richette P, Clerson P, Perissin L, Flipo RM, Bardin T. Revisiting comorbidities in gout: a cluster analysis. *Ann Rheum Dis* 2015;74:142–7.
- Newcomer SR, Steiner JF, Bayliss EA. Identifying subgroups of complex patients with cluster analysis. *Am J Manag Care* 2011;17:e324–32.
- Carpenter L, Nikiphorou E, Kiely PDW, Walsh DA, Young A, Norton S. Secular changes in the progression of clinical markers and patient-reported outcomes in early rheumatoid arthritis. *Rheumatology (Oxford)* 2020;59:2381–91.
- Carpenter L, Barnett R, Mahendran P, Nikiphorou E, Gwinnutt J, Verstappen S, et al. Secular changes in functional disability, pain, fatigue and mental well-being in early rheumatoid arthritis. A longitudinal meta-analysis. *Semin Arthritis Rheum* 2020;50:209–19.
- Geenen R, Overman CL, Christensen R, Asenlof P, Capela S, Huisinga KL, et al. EULAR recommendations for the health professional's approach to pain management in inflammatory arthritis and osteoarthritis. *Ann Rheum Dis* 2018;77:797–807.
- Gron KL, Ornbjerg LM, Hetland ML, Aslam F, Khan NA, Jacobs JW, et al. The association of fatigue, comorbidity burden, disease activity, disability and gross domestic product in patients with rheumatoid arthritis. Results from 34 countries participating in the Quest-RA program. *Clin Exp Rheumatol* 2014;32:869–77.
- Lee YC, Frits ML, Iannaccone CK, Weinblatt ME, Shadick NA, Williams DA, et al. Subgrouping of patients with rheumatoid arthritis based on pain, fatigue, inflammation, and psychosocial factors. *Arthritis Rheumatol* 2014;66:2006–14.
- Ziade N, El Khoury B, Zoghbi M, Merheb G, Abi Karam G, Mroue K, et al. Prevalence and pattern of comorbidities in chronic rheumatic and musculoskeletal diseases: the COMORD study. *Sci Rep* 2020;10:7683.
- [Update on current care guidelines. Rheumatoid Arthritis (RA)]. *Duodecim* 2015;131:1409–10.
- Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther* 2009;11:R7.
- Aktas A, Walsh D, Rybicki L. Symptom clusters: myth or reality? *Palliat Med* 2010;24:373–85.

## Supporting information

Additional Supporting Information may be found in the online version of this article.

[Supplementary Figure 1](#)

[Supplementary Figure 2](#)

[Supplementary Figure 3](#)

[Supplementary Figure 4](#)

[Supplementary Figure 5](#)

[Supplementary Figure 6](#)

[Supplementary Figure 7](#)

[Supplementary Table 1](#)

[Supplementary Table 2](#)

[Supplementary Table 3](#)

[Supplementary Table 4](#)

Please note that the editors are not responsible for the content or functionality of any supplementary material supplied by the authors. Any queries should be directed to the corresponding author.