**Exploring the disparity between inflammation and disability in the ten year outcomes of people with rheumatoid arthritis**

James M Gwinnutt[1] – james.gwinnutt@manchester.ac.uk (ORCID: 0000-0002-1435-8797)

Sam Norton[2,3] – sam.norton@kcl.ac.uk

Kimme L Hyrich[1,4] – Kimme.hyrich@manchester.ac.uk (ORCID: 0000-0001-8242-9262)

Mark Lunt[1] – mark.lunt@manchester.ac.uk (ORCID: 0000-0002-2391-5575)

Bernard Combe[5] – bernard.combe@umontpellier.fr

Nathalie Rincheval[6] – nathalie.rincheval@inserm.fr

Adeline Ruyssen-Witrand[7, 8] - ruyssen-witrand.a@chu-toulouse.fr

Bruno Fautrel[9, 10] - bruno.fautrel@aphp.fr

Daniel F McWilliams[11, 12] – dan.mcwilliams@nottingham.ac.uk (ORCID 0000-0002-0581-1895)

David A Walsh[11, 12, 13] – david.walsh@nottingham.ac.uk

Elena Nikiphorou[3, 14] – enikiphorou@gmail.com

Patrick Kiely[15, 16] - patrick.kiely@nhs.net

Adam Young[17] - adam.young@nhs.net

Jacqueline R Chipping[18,19] – j.chipping@uea.ac.uk

Alex MacGregor[18,19] – a.macgregor@uea.ac.uk

Suzanne MM Verstappen[1,4] – Suzanne.verstappen@manchester.ac.uk (ORCID: 0000-0001-6181-0646)

[1] Centre for Epidemiology Versus Arthritis, Division of Musculoskeletal and Dermatological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK
[2] Health Psychology section, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[3] Centre for Rheumatic Diseases, Department of Inflammation Biology, Faculty of Life Sciences and Medicine, King's College London, London, UK
[4] NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, UK
[5] University of Montpellier, Montpellier, France
[6] Laboratory of Biostatistics and Epidemiology, University of Montpellier, Montpellier, France
[7] Centre de Rhumatologie, Hôpital Purpan, Toulouse, France
[8] Faculté de Médecine, Université Toulouse III, Paul Sabatier University, Inserm UMR1027, Toulouse, France
[9] Sorbonne University – Assistance Publique Hôpitaux de Paris, Pitie Salpetriere Hospital, Department of Rheumatology, Paris, France
[10] PEPITES team, Pierre Louis Institute of Epidemiology and Public Health, INSERM UMRS 1136, Paris France
[11] Pain Centre Versus Arthritis, University of Nottingham, Nottingham, UK
[12] NIHR Nottingham Biomedical Research Centre, Nottingham, UK
[13] Department of Rheumatology, Sherwood Forest Hospitals NHS Foundation Trust, Sutton in Ashfield, UK
[14] Rheumatology Department, King's College Hospital, London, UK
[15] Department of Rheumatology, St George's University Hospitals NHS Foundation Trust, UK
[16] Institute of Medical and Biomedical Education, St George's University of London, London, UK

[17] Centre for Health Services and Clinical Research, Life and Medical Sciences, University of Hertfordshire, Hatfield, UK
[18] Norwich Medical School, University of East Anglia, Norwich, UK
[19] Rheumatology Department, Norfolk and Norwich University Hospitals NHS Trust, Norwich, UK

Word count: 3273 (max: 3500)

Corresponding author: James M Gwinnutt, Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. James.gwinnutt@manchester.ac.uk

Key words: rheumatoid arthritis, disability, epidemiology, outcomes research, psychology

## Abstract

**Objectives:** To identify groups of people with rheumatoid arthritis (RA) with different disability trajectories over ten years, despite comparable levels of inflammation.

**Methods:** Data for this analysis came from three European prospective cohort studies of people with RA (Norfolk Arthritis Register [NOAR], Early Rheumatoid Arthritis Network [ERAN], Étude et Suivi des Polyarthrites Indifférenciées Récentes [ESPOIR]). Participants were assessed regularly over 8 (ERAN) to 10 (NOAR/ESPOIR) years. Inclusion criteria were: recruited after 1/1/2000, <24 months baseline symptom duration, and disability (Health Assessment Questionnaire [HAQ]) and inflammation (two-component disease activity score [DAS28-2C]) recorded at baseline and one other follow-up. People in each cohort also completed patient reported outcome measures at each assessment (pain, fatigue, depression). Group-based trajectory models (GBTM) were used to identify distinct groups of people with similar HAQ and DAS28-2C trajectories over follow-up.

**Results:** This analysis included 2500 people with RA (NOAR: 1000, ESPOIR: 766, ERAN: 734). People in ESPOIR were younger and included more women (mean [standard deviation] age: NOAR: 57.1 [14.6], ESPOIR: 47.6 [12.5], ERAN: 56.8 [13.8]; women: NOAR: 63.9%, ESPOIR: 76.9%, ERAN: 69.1%). Within each cohort, two pairs of trajectories that followed the hypothesised pattern (comparable DAS28-2C but different HAQ) were identified. Higher pain, fatigue and depression were associated with increased odds of being in the high HAQ trajectories.

**Conclusion:** Excess disability is persistent in RA. Controlling inflammation may not be sufficient to alleviate disability in all people with RA, and effective pain, fatigue and mood management may be needed in some groups to improve long-term function.

Words: 250 (max 250)

Rheumatoid arthritis (RA) is a condition characterised by inflammation of synovial joints.(1) In the past, limited treatment options were available to control this inflammation and therefore people with RA suffered from significant pain and disability into the long-term.(2) However, following the adoption of treat-to-target strategies and the widespread use of methotrexate for RA in the mid-1990s and subsequently the introduction of biologic treatments in the 2000s,(3, 4) the ability to control inflammation drastically improved, leading to low inflammation over time for many people with RA.(5, 6)

Nonetheless, this low long-term inflammation has not translated into low levels of disability. A study from the Norfolk Arthritis Register (NOAR) showed that disability followed a "J-shaped" trajectory over ten-years, culminating in disability levels similar to baseline.(7) The same trajectory has been observed in other UK,(6) Swedish,(8) and French cohorts,(9) and within a longitudinal meta-analysis.(10) Furthermore, this disparity between inflammation and disability was larger in the 2000s than in the 1990s,(7, 10) despite the increasing options available to control inflammation.

Disability impacts all aspects of the lives of people with RA; higher disability is associated with reductions in work capacity(11, 12) and interference with valued life activities such as seeing friends and taking care of family.(13) Furthermore, disability is potentially a significant burden for healthcare systems. Disability is the strongest predictor of healthcare costs in RA, a finding seen across several healthcare settings.(14-16) Therefore, this excess disability despite treatment of inflammation requires investigation.

Not all individuals follow the same symptom trajectory, and the progression of many long-term outcomes important to people with RA can be described using multiple sub-groups or "trajectory groups".(17-19) The hypothesis of this research project is that the disparity between inflammation and disability seen on average in cohorts of people with RA is driven by a sub-group of people with RA characterised by low-inflammation yet high disability into the long-term. The aim of this analysis was to identify this subgroup within three large-scale cohort studies of people with inflammatory arthritis, two from the United Kingdom and one from France. Then, we aimed to identify factors driving the excess disability in this subgroup.

**Methods**

The data for this analysis came from three inception cohorts of people with inflammatory arthritis (IA). Participants in all three studies provided written informed consent. The Norfolk Arthritis Register (NOAR) is a primary-care based, prospective inception cohort of people with IA recruited in Norfolk, UK.(20) The inclusion criteria for NOAR are ≥2 swollen joints lasting for ≥4 weeks and being ≥16 years old. Recruitment started in 1990 and is ongoing. Participants in NOAR were assessed at baseline and then at years 1, 2, 3, 5, 7 and 10. NOAR was approved by the Cambridgeshire and Hertfordshire Research Ethics Committee (15/EE/0076).

The Early Rheumatoid Arthritis Network (ERAN) is a cohort of people recruited at the point of clinician diagnosis of RA from 22 outpatient rheumatology clinics in the UK and Ireland from 2002 to 2013.(21) Participants of ERAN were seen at baseline, once between 3 and 6 months, and then annually thereafter for up to 13 years. Only the first eight years of follow-up within ERAN were used for the current analysis due to attrition, largely driven by all centres closing to follow-up by 2018. ERAN was approved by the Trent Research Ethics Committee (01/4/047).

The *Étude et Suivi des Polyarthrites Indifférenciées Récentes* (ESPOIR) study is a cohort of people with IA recruited from 14 regional centres of rheumatology across France between 2002 and 2005. The inclusion criteria for ESPOIR were >2 swollen joints lasting for >6 weeks, clinical diagnosis of RA as certain or possible, and being aged 18-70 years. Furthermore, participants were required to have received no DMARDs or glucocorticoids for >2 weeks.(22) Participants were assessed at baseline, 6 months, 12 months, 18 months, 24 months and then yearly up to 10 years. The ESPOIR cohort study was approved by the Ethics Committee of Montpellier (020307).

For the current analysis, participants within each cohort were included if they had <24 months symptom duration at baseline, had data for disability and inflammation at baseline and one other assessment, and were recruited on or after the year 2000.

*Assessments*

Participants in each study reported demographics (age, gender, smoking status, symptom duration) and completed questionnaires. Research nurses measured height and weight and performed swollen and tender joint counts at each assessment. Blood samples were taken at each assessment from which C-reactive protein (CRP) was measured in NOAR and ESPOIR, and erythrocyte sedimentation rate (ESR) in ERAN. Rheumatoid factor (RF; all cohorts) and anti-cyclic citrullinated protein antibody (anti-CCP; NOAR and ESPOIR only) positivity were measured from baseline blood samples. Prescription disease modifying anti-rheumatic drug (DMARD) treatments were also recorded. The ERAN dataset includes the rheumatic disease comorbidity index (RDCI),(23) whereas participants of NOAR and ESPOIR self-reported comorbidities from predetermined lists (coded as 0, 1 or ≥2 comorbidities due to insufficient data to calculate RDCI). Data on baseline joint erosions were available in the ERAN and ESPOIR cohorts. X-rays were not routinely taken as part of the NOAR assessments (74.2% missing X-rays as per NOAR protocol, see Supplementary Table 1).

Global disease activity measures (such as the Disease Activity Score 28 [DAS28]) include both inflammatory markers (swollen joint count, CRP/ESR) and patient reported outcome measures (PROMS; tender joint counts, global health visual analogue scale [VAS]). However, pharmacological treatment of RA aims to reduce inflammation and previous research has shown that only the inflammatory components of the DAS28 were associated with MRI-detected synovitis.(24) Furthermore, this study aimed to identify specific factors driving disability in RA, such as inflammation and PROMs. As these factors are conflated in global disease activity measures, a measure of inflammation alone was needed. Therefore, in this analysis inflammation was quantified using the two-component DAS28 (DAS28-2C). This measure combines swollen joint count and either CRP (NOAR, ESPOIR) or ESR (ERAN) using formulae designed to maximise the association between the scores and ultrasound synovitis.(25) Participants also completed several PROMs. Disability was assessed using the British(26) (NOAR, ERAN) and French(27) (ESPOIR) versions of the Health Assessment Questionnaire (HAQ), with scores adjusted to account for device use. Participants of NOAR and ESPOIR completed pain and fatigue VAS, and participants of ERAN completed the Short-Form 36 (SF-36), which includes pain and fatigue (vitality) subscales.(28) Anxiety and depression were assessed using the Arthritis Impact Measurement Scale (AIMS-2)(29) in NOAR, five variables from the French version of the AIMS-2(30) in ESPOIR, and the mental health component of the SF36 in ERAN.

*Statistical analysis*

Demographic and clinical characteristics were summarised using descriptive statistics, stratified by cohort. Within each cohort, subgroups of participants with similar HAQ and DAS28-2C trajectories were identified using multivariate group-based trajectory analysis,(31) a longitudinal finite-mixture

model. Specifically, this involved jointly estimating longitudinal models for both HAQ and DAS28-2C that estimated the baseline level of the outcome (intercept) and the rate of change in the outcome over time (slope), with these variables used to inform the identification of trajectory subgroups (latent classes). The number of trajectory groups was selected by assessing the Akaike and Bayesian Information Criteria, entropy and posterior probability of group membership (Supplementary Document for details). Within pairs of trajectories that displayed the hypothesised relationship (similar inflammation but different HAQ trajectories over follow-up), baseline predictors of being in the group characterised by higher HAQ score trajectory were assessed using multivariable logistic regression. Missing data on baseline predictors were imputed using multiple imputation by iterative chained equations. Outcomes over eight (ERAN) and ten (NOAR and ESPOIR) years of follow-up were compared between trajectory groups using linear mixed models for continuous outcomes and generalised estimating equations (GEE) analysis for binary outcomes, controlling for age and gender. The associations between time-varying PROMs (pain, fatigue, anxiety and depression) and disability were assessed using mixed effects models, controlling for age, gender, baseline comorbidity, baseline BMI, and HAQ score at the previous assessment. As the PROMS were measured on different scales (e.g. VAS pain [0-100] and AIMS depression [0-10]), to improve comparability, the PROMS were also standardised (i.e. rescaled to have mean of 0 and standard deviation of 1). Using standardised PROMS, model coefficients represent a change in the HAQ score for a standard deviation change in the PROMS (results in supplementary file). Interaction terms were included to assess whether the association between PROMs and disability differed between trajectory groups. Trajectory analysis was performed using the traj package(32) in Stata version 14 (StataCorp: College Station, TX), and other analyses performed using R version 3.6.0 (packages: haven,(33) tidyverse,(34) grid, gridExtra,(35) reshape2,(36) lme4,(37) psych,(38) mice,(39) miceadds,(40) effects,(41) gee,(42) broom.mixed(43)).

**Results**

This analysis included 2500 people with inflammatory arthritis (NOAR = 1000, ESPOIR = 766, ERAN = 734). The ESPOIR participants were younger than the NOAR and ERAN participants (mean age [standard deviation], years: ESPOIR 47.6 [12.5]; NOAR 57.1 [14.6]; ERAN 56.8 [13.8]) and had a higher proportion of women (% women: ESPOIR 76.9%; NOAR 63.9%; ERAN 69.1%). The ESPOIR participants had shorter symptom duration, had more severe disease, and fewer participants were receiving csDMARDs at baseline compared with NOAR and ERAN (Supplementary Table 1).

*Group-based trajectory analysis*

Assessment of group-based trajectory models applied to the longitudinal HAQ and DAS28-2C scores in each cohort separately resulted in the selection of a five-group trajectory model (see Supplementary Table 2 and Supplementary Figures 1-4). Each cohort contained one trajectory group with very low HAQ and DAS28-2C scores (group 1 in Figure 1 [yellow trajectory], termed “Very low inflammation-Low HAQ”) (Supplementary Table 3 for baseline characteristics). In this group, the HAQ and DAS28-2C scores remained low over follow-up. The hypothesised relationship (similar inflammation [DAS28-2C] but different disability [HAQ] trajectories) was observed in two pairs of trajectories in each cohort (Figure 1). Within each pair, the DAS28-2C scores were similar but one trajectory had an average HAQ score of 0.5-1.0 unit higher over follow-up (groups 3 and 5 in Figure 1 [dashed lines], termed “high HAQ” trajectories) than the other trajectory (groups 2 and 4 in Figure 1 [solid lines], termed “low HAQ” trajectories) over 8-10 years. In each cohort, one pair of trajectories had lower disability and inflammation on average over the course of follow-up (groups 2 and 3 in Figure 1 [purple trajectories], termed “Low inflammation pair”) compared with the other pair (groups 4 and 5 in Figure 1 [green trajectories], termed “High inflammation pair”). In general, the inflammation scores of these trajectory groups improved over follow-up, whereas the HAQ scores were relatively

stable. In summary, the five trajectory groups were: 1 = “Very low inflammation-Low HAQ” (NOAR: 28.7%; ESPOIR: 24.3%; ERAN: 14.9%), 2 = “Low inflammation-Low HAQ” (NOAR: 29.5%; ESPOIR: 29.8%; ERAN: 11.9%), 3 = “Low inflammation-High HAQ” (NOAR: 19.9%; ESPOIR: 16.6%; ERAN: 28.3%), 4 = “High inflammation-Low HAQ” (NOAR: 10.4%; ESPOIR: 17.8%; ERAN: 28.7%), and 5 = “High inflammation-High HAQ” (NOAR: 11.5%; ESPOIR: 11.6%; ERAN: 16.2%).

*Baseline factors associated with high HAQ trajectory group membership*

At baseline, participants in the “Low inflammation-High HAQ” group were on average older, were more often women, had more comorbidities and had more severe pain, fatigue, anxiety and depression compared with the “Low inflammation-Low HAQ” group, despite similar inflammation (Table 1). Similar results were seen when comparing the “High inflammation-High HAQ” group with the “High inflammation-Low HAQ” group. Furthermore, in the low inflammation pair, the high HAQ trajectory had more erosions at baseline in ERAN and ESPOIR (data not available in NOAR). This was not seen in the high inflammation pair.

Multivariable logistic regression analysis was used to identify baseline factors associated with high HAQ trajectory membership compared with low HAQ trajectory membership. Separate models were constructed for the high and low inflammation pairs (i.e. the “Low inflammation-High HAQ” group was compared with the “Low inflammation-Low HAQ group” and the “High inflammation-High HAQ” group was compared with the “High inflammation-Low HAQ group”). Older age, being a woman vs. a man, and more severe pain, fatigue and depression were associated with increased odds of being in the higher HAQ trajectory in both the high and low inflammation pairs (Table 2; see Supplementary Table 4 for sensitivity analysis regarding missing data in NOAR). More comorbidities and serology status (NOAR: anti-CCP+; ERAN: RF+) were associated with greater odds of being in the high HAQ trajectories in NOAR and ERAN, although with wide confidence intervals which included the null for comorbidities. Erosions were associated with being in the high HAQ trajectory in the low inflammation pair in ERAN and ESPOIR, but not in the high inflammation pair, although the estimates were imprecise.

*Outcomes over time*

The high HAQ trajectories had greater tender joint counts, pain, fatigue, depression and anxiety than the low HAQ trajectories over follow-up in both inflammation pairs (Table 3). The high HAQ trajectories also had more comorbidities over time compared with the low HAQ trajectories across the cohorts, but with wide confidence intervals containing the null (Table 3).

Across all trajectory groups, more severe scores on PROMs (pain, fatigue, anxiety and depression) were all associated with increasing HAQ scores measured at the same assessment (Table 4, Supplementary Table 5 for unimputed analysis), independent of age, gender, baseline comorbidity and BMI, and HAQ at the previous assessments (see Supplementary Figure 5 for directed acyclic graph underpinning this analysis). After standardising the PROMs to improve comparability, pain had the strongest association with HAQ (Supplementary Table 6). However, the relationship between PROMs and HAQ was different between the high and low HAQ trajectories (Figure 2). Particularly in the high inflammation pairs, the association between the PROMs and HAQ score was stronger (i.e. the slope was steeper [interaction terms in Table 4]) in the low HAQ trajectory compared with the high HAQ trajectory (Figure 2).

**Discussion**

This large-scale analysis of 2500 people with RA with follow-up of 8-10 years illustrates the disparity between inflammation and disability for many people with this disease. This analysis identified two pairs of trajectories within each cohort, one classified as “high inflammation” (groups 4 & 5) and the other classified as “low inflammation” (groups 2 & 3). Each trajectory pair is characterised by a high and a low disability trajectory, despite similar inflammation scores. Therefore, 30-45% of people with RA were in groups characterised by potentially excess disability (i.e. groups 3 & 5: NOAR: 31.4%, ESPOIR: 28.2%, ERAN: 44.5%), over and above what would be expected from their level of inflammation. This excess disability is maintained up to a decade following onset. People in the higher disability trajectories were older, were more likely to be women and had worse pain, fatigue and mental health compared with the lower disability trajectories. In the high HAQ trajectories, the relationship between PROMs and disability was weaker compared with the low HAQ trajectories, particularly in the high inflammation pairs.

Our findings extend results from previous studies. A trajectory analysis of 9493 people with RA reported that 65% of those with controlled inflammation still reported persistent pain over three years.(18) A cohort of 232 people with early RA reported that 34% of participants had unacceptable pain at 5 years and this was associated with lower inflammation at baseline.(44) Van der Elst et al. demonstrated that one in five people with rapidly and persistently controlled early RA still reported high fatigue and/or pain after one year of follow-up.(45) A cross-sectional analysis of 169 people with RA reported a subgroup of 57 people (33.7%) who had high pain, fatigue and depression despite low inflammation scores.(46) In summary, for many people with RA, controlling inflammation is not sufficient to alleviate symptom burden. Our analysis illustrates these findings are consistent across several cohorts, and persist for up to a decade, with no group showing improvements in disability commencing later in follow-up, indicating once more the importance of early intervention in RA.

As PROMs (e.g. pain, fatigue, mental health) are the most consistent predictors of long-term function,(47) we investigated the relationship between these PROMs and disability, both when measured at baseline and longitudinally. Baseline and time-varying pain, fatigue and mental health were all associated with high HAQ trajectory membership. In 2006, Aletaha et al. described reversible and irreversible components of disability in RA,(48) whereby the reversible component of disability is driven by current inflammation and the irreversible component driven by joint damage and co-existing conditions. This analysis demonstrates the potential impact of comorbidities and erosion, showing that baseline erosions predicted high HAQ trajectory membership in ESPOIR and ERAN and that the high HAQ trajectories had more comorbid conditions compared with the low HAQ trajectories.(49, 50) Erosions at baseline could be as a result of treatment delays, further emphasising the importance of early treatment in RA. Furthermore, the current analysis suggests that a third component of RA disability may comprise pain, fatigue and mental health, given the large differences in disability between pairs of trajectory groups with similar inflammation yet large differences in these PROMs. Whereas pain, fatigue and poor mental health may not be as irreversible as joint erosion, they are challenging to ameliorate in people with RA and may require both pharmacological and non-pharmacological interventions not targeting inflammation.(51-53)

Despite the higher PROM scores in the high HAQ trajectories compared with the lower HAQ trajectories, a surprising result was the interaction between trajectory membership and PROMs when predicting disability. This analysis reported a weaker association between PROMs and disability in the high HAQ trajectories compared with the low HAQ trajectories. This observation could be due to the ceiling effect of the HAQ score,(54) particularly as this was primarily seen in the high inflammation pair with HAQ scores nearer the top of the scale.

The strengths of this analysis include the large sample size and long-term follow-up, meaning that the disparity between inflammation and disability in people with RA over ten years could be precisely characterised for the first time. The cohorts were well phenotyped, meaning a large array of potential factors driving disability could be assessed. Whilst the inclusion criteria were similar between the three cohorts, there were several differences between the demographic and clinical characteristics of the cohorts at baseline, due to when participants were recruited (with ESPOIR's participants recruited earlier in the disease process compared with NOAR and ERAN). The participants of ESPOIR were on average 8-10 years younger than the NOAR and ERAN cohorts. This could in part be explained by the higher rate of smoking in the ESPOIR population and the inclusion criterion of being less than 70 years old at baseline. Whilst age was controlled for in the regression analyses of this project, the large difference in age between the cohorts may still be affecting the results. There were also a number of differences in terms of the PROMs included in the three studies (i.e. the measures of pain, fatigue, mental health). A further limitation of this study was the use of different blood sample analyses across the cohorts, with ESR used in ERAN and CRP in NOAR and ESPOIR when calculating the DAS28-2C. As yet, the comparability of the DAS28-2C-ESR and DAS28-2C-CRP has not been established. Despite these limitations, the consistent results across the three cohorts suggest generalisability of the findings. Whilst these cohorts had large samples, some effect estimates had wide confidence intervals and overlapped the null, and therefore some caution should be taken when interpreting these estimates. There was a significant proportion of missing data in some of the PROMs in the NOAR cohort, which clustered in the participants recruited earlier in the cohort and could result in bias in the reported associations. Multiple imputation was used to impute missing data and further sensitivity analysis indicated minimal missing data bias (Supplementary Table 7). The names of the trajectory groups (i.e. high / low inflammation) are used in relation to one another, rather than based on external definitions of high and low inflammation. However, as these cohorts are representative samples of the population of people with RA, these results indicate that there is excess disability evident in RA across the spectrum of inflammation levels seen in clinical care. The scales to measure mental health across the three cohorts were necessarily brief, given the large amount of data collected at each assessment. However, they do not provide an unambiguous measurement of depression or anxiety. Lastly, there was significant attrition over follow-up. This was particularly substantial in ERAN, meaning only eight years of follow-up could be included.

In conclusion, this analysis illustrates that approximately 30-45% of people with RA have excess disability (i.e. discordant with inflammation level), and that this excess disability is seen across inflammation levels. This excess disability is persistent, with disparity remaining at least up to 10 years following onset. People with RA in the high disability trajectories had more severe pain, fatigue and depression compared with those in low disability trajectories, despite similar inflammation levels. This indicates the urgent need to address pain, fatigue and depression, for example by psychological interventions for people with RA, in order to curtail long-term disability.

**KEY MESSAGES (15 words max)**

- Previous research shows a disparity between long-term low inflammation and high disability in rheumatoid arthritis
- This study identified groups with similar inflammation trajectories yet markedly different disability over ten years
- Pain, fatigue and depression predicted higher disability group membership, independent of inflammation

**ACKNOWLEDGEMENTS**

The authors thank the participants involved in NOAR, ERAN and ESPOIR, as well as the clinical staff at each of the recruiting centres. Thanks also to the data management team at the University of Manchester, University of Nottingham and University of Montpellier. This research was supported by the NIHR Manchester Biomedical Research Centre and the NIHR Nottingham Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

ESPOIR cohort: An unrestricted grant from Merck Sharp and Dohme (MSD) was allocated for the first 5 years. Two additional grants from INSERM were obtained to support part of the biological database. The French Society of Rheumatology, Abbvie, Pfizer, Lilly, and more recently Fresenius and Galapagos also supported the ESPOIR cohort study.

ERAN cohort: Unrestricted grants from Wyeth Pharmaceuticals and the UK Healthcare Commission supported creation of the ERAN database

**FUNDING**

This work was supported by the Medical Research Council (through a Skills Development Fellowship for JMG), Versus Arthritis (grant numbers 20385, 20380) and supported by the NIHR Manchester Biomedical Research Centre and the NIHR Nottingham Biomedical Research Centre.

**CONTRIBUTORS**

Review of manuscript: JMG, SN, KLH, ML, BC, NR, ARW, BF, DFM, DAW, EN, AY, PK, JRC, AM, SMMV; Study concept and design: JMG, SN, KLH, SMMV; Acquisition of data: BC, NR, ARW, BF, DFM, DAW, EN, AY, PK, JRC, AM, SMMV; Analysis and interpretation of data: JMG, SN, KLH, ML, SMMV

**COMPETING INTERESTS**

There were no competing interests directly relevant to this manuscript. Other disclosures include: JMG: grants (Bristol-Myers Squibb); SN: None; KLH: grants (Bristol-Myers Squibb, Pfizer), honoraria (Abbvie); ML: None; BC: grants (Novartis, Pfizer, Roche), honoraria (Abbvie, Bristol-Myers Squibb, Gilead-Galapagos, Janssen, Lilly, Merck, Novartis, Pfizer, Roche-Chugai); NR: None; ARW: Consulting (Pfizer, Abbvie, Novartis, Lilly, Janssen), honoraria (Abbvie, Bristol-Myers Squibb, Galapagos, Fresenius-Kabi, Mylan-Viatris, MSD, Novartis, Lilly, UCB, Pfizer, Roche-Chugaï, Sanofi), support for attending meetings (Abbvie, Amgen, Fresenius-Kabi); BF: None; DFM: grants (Pfizer, Eli Lilly); DAW: grants (Pfizer, Eli Lilly), consulting (Pfizer, Abbvie, GSK, Reckitt-Benckiser, Galapagos, Eli Lilly), honoraria (Pfizer, Abbvie); EN: grants (Pfizer, Lilly), honoraria (Celltrion, Pfizer, Sanofi, Gilead, Galapagos, Abbvie, Lilly); PDWK: honoraria (Lilly, Galapagos, Vifor); AY: None; JRC: None; AM: None; SMMV: None.

**DATA ACCESS STATEMENT**

The data underlying this article cannot be shared publicly for the privacy of individuals that participated in the study. The data will be shared on reasonable request to the principal investigators of each of the three datasets (NOAR: AM, ERAN: DAW, ESPOIR: BC).

**REFERENCES**

1. Smolen JS, Aletaha D, Barton A, Burmester GR, Emery P, Firestein GS, et al. Rheumatoid arthritis. Nat Rev Dis Primers. 2018;4:18001.

2. Scott DL, Symmons DP, Coulton BL, Popert AJ. Long-term outcome of treating rheumatoid arthritis: results after 20 years. Lancet. 1987;1(8542):1108-11.
3. Pincus T, Yazici Y, Sokka T, Aletaha D, Smolen JS. Methotrexate as the "anchor drug" for the treatment of early rheumatoid arthritis. Clin Exp Rheumatol. 2003;21(5 Suppl 31):S179-S85.
4. Olsen NJ, Stein CM. New drugs for rheumatoid arthritis. N Engl J Med. 2004;350(21):2167-79.
5. Gwinnutt JM, Symmons DPM, Macgregor AJ, Chipping JR, Marshall T, Lunt M, et al. Twenty-Year Outcome and Association Between Early Treatment and Mortality and Disability in an Inception Cohort of Patients With Rheumatoid Arthritis: Results From the Norfolk Arthritis Register. Arthritis Rheumatol. 2017;69(8):1566-75.
6. 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, England). 2020:kez635.
7. Gwinnutt JM, Symmons DPM, Macgregor AJ, Chipping JR, Marshall T, Lunt M, et al. Have the 10-year outcomes of patients with early inflammatory arthritis improved in the new millennium compared with the decade before? Results from the Norfolk Arthritis Register. Ann Rheum Dis. 2018;77(6):848-54.
8. Kapetanovic MC, Lindqvist E, Nilsson JA, Geborek P, Saxne T, Eberhardt K. Development of functional impairment and disability in rheumatoid arthritis patients followed for 20 years: relation to disease activity, joint damage, and comorbidity. Arthritis Care Res (Hoboken ). 2015;67(3):340-8.
9. Courvoisier N, Dougados M, Cantagrel A, Goupille P, Meyer O, Sibilia J, et al. Prognostic factors of 10-year radiographic outcome in early rheumatoid arthritis: a prospective study. Arthritis Res Ther. 2008;10(5):R106.
10. 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. 2019:S0049-172(19)30097-6.
11. Boot CRL, de Wind A, van Vilsteren M, van der Beek AJ, van Schaardenburg D, Anema JR. One-year Predictors of Presenteeism in Workers with Rheumatoid Arthritis: Disease-related Factors and Characteristics of General Health and Work. J Rheumatol. 2018;45(6):766-70.
12. Gwinnutt JM, Leggett S, Lunt M, Barton A, Hyrich KL, Walker-Bone K, et al. Predictors of presenteeism, absenteeism and job loss in patients commencing methotrexate or biologic therapy for rheumatoid arthritis. Rheumatology (Oxford). 2020;59(10):2908-19.
13. Katz PP, Morris A, Yelin EH. Prevalence and predictors of disability in valued life activities among individuals with rheumatoid arthritis. Ann Rheum Dis. 2006;65(6):763-9.
14. Wallman JK, Eriksson JK, Nilsson JA, Olofsson T, Kristensen LE, Neovius M, et al. Costs in Relation to Disability, Disease Activity, and Health-related Quality of Life in Rheumatoid Arthritis: Observational Data from Southern Sweden. J Rheumatol. 2016;43(7):1292-9.
15. Ohinmaa AE, Thanh NX, Barnabe C, Martin L, Russell AS, Barr SG, et al. Canadian estimates of health care utilization costs for rheumatoid arthritis patients with and without therapy with biologic agents. Arthritis Care Res (Hoboken ). 2014;66(9):1319-27.
16. Michaud K, Messer J, Choi HK, Wolfe F. Direct medical costs and their predictors in patients with rheumatoid arthritis: a three-year study of 7,527 patients. Arthritis Rheum. 2003;48(10):2750-62.
17. Norton S, Sacker A, Dixey J, Done J, Williams P, Young A. Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. Rheumatology (Oxford). 2013;52(11):2016-24.
18. McWilliams DF, Dawson O, Young A, Kiely PDW, Ferguson E, Walsh DA. Discrete Trajectories of Resolving and Persistent Pain in People With Rheumatoid Arthritis Despite Undergoing Treatment for Inflammation: Results From Three UK Cohorts. The journal of pain : official journal of the American Pain Society. 2019;20(6):716-27.

19. Druce KL, Jones GT, Macfarlane GJ, Verstappen SM, Basu N. The Longitudinal Course of Fatigue in Rheumatoid Arthritis: Results from the Norfolk Arthritis Register. J Rheumatol. 2015;42(11):2059-65.
20. Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. Br J Rheumatol. 1994;33(8):735-9.
21. Garwood W. The Early Rheumatoid Arthritis Network (ERAN). Musculoskeletal Care. 2004;2(4):240-4.
22. Combe B, Benessiano J, Berenbaum F, Cantagrel A, Daurès JP, Dougados M, et al. The ESPOIR cohort: a ten-year follow-up of early arthritis in France: methodology and baseline characteristics of the 813 included patients. Joint Bone Spine. 2007;74(5):440-5.
23. England BR, Sayles H, Mikuls TR, Johnson DS, Michaud K. Validation of the rheumatic disease comorbidity index. Arthritis Care Res (Hoboken). 2015;67(6):865-72.
24. Baker JF, Conaghan PG, Smolen JS, Aletaha D, Shults J, Emery P, et al. Development and validation of modified disease activity scores in rheumatoid arthritis: superior correlation with magnetic resonance imaging-detected synovitis and radiographic progression. Arthritis Rheumatol. 2014;66(4):794-802.
25. Hensor EMA, McKeigue P, Ling SF, Colombo M, Barrett JH, Nam JL, et al. Validity of a two-component imaging-derived disease activity score for improved assessment of synovitis in early rheumatoid arthritis. Rheumatology (Oxford). 2019.
26. Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol. 1986;25(2):206-9.
27. Guillemin F, Braincon S, Pourel J. [Measurement of the functional capacity in rheumatoid polyarthritis: a French adaptation of the Health Assessment Questionnaire (HAQ)]. Rev Rhum Mal Osteoartic. 1991;58(6):459-65.
28. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. J Epidemiol Community Health. 1999;53(1):46-50.
29. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. Arthritis Rheum. 1992;35(1):1-10.
30. Pouchot J, Guillemin F, Coste J, Brégeon C, Sany J. Validity, reliability, and sensitivity to change of a French version of the arthritis impact measurement scales 2 (AIMS2) in patients with rheumatoid arthritis treated with methotrexate. J Rheumatol. 1996;23(1):52-60.
31. Jones B, Nagin DS. Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them. Sociological Methods and Research. 2007;35(4):542-71.
32. Jones B, Nagin D. A note on a Stata plugin for estimating group-based trajectory models. Sociological Methods & Research. 2013;42(4):608-13.
33. Wickham H, Miller E. haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.3.1. https://CRAN.R-project.org/package=haven 2020 [Available from: https://CRAN.R-project.org/package=haven.
34. Wickham H, Averick M, Bryan J, Chang W, D'Agostino McGowan L, François R, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4(43):1686.
35. Auguie B. gridExtra: Miscellaneous Functions for "Grid" Graphics 2017 [Available from: https://CRAN.R-project.org/package=gridExtra.
36. Wickham H. Reshaping Data with the reshape Package. Journal of Statistical Software. 2007;21(12):20.
37. Bates D, Maechler M, Bolker M, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 2015;67(1):1-48.
38. Revelle W. psych: Procedures for Personality and Psychological Research 2020 [15.6.21]. Available from: https://CRAN.R-project.org/package=psych.

39. van Buuren S, Groothuis-Oudshoorn L. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011;45(3):1-67.
40. Robitzsch A, Grund S. miceadds: Some Additional Multiple Imputation Functions, Especially for 'mice' 2020 [Available from: https://CRAN.R-project.org/package=miceadds.
41. Fox J, Weisberg S. An R Companion to Applied Regression (3rd Edition). Thousand Oaks, CA: SAGE; 2019.
42. Carey V, Lumley T, Ripley B. gee: Generalized Estimation Equation Solver 2019 [Available from: https://CRAN.R-project.org/package=gee.
43. Bolker B, Robinson D. broom.mixed: Tidying Methods for Mixed Models. 2020 [Available from: https://CRAN.R-project.org/package=broom.mixed.
44. Eberhard A, Bergman S, Mandl T, Olofsson T, Rydholm M, Jacobsson L, et al. Predictors of unacceptable pain with and without low inflammation over 5 years in early rheumatoid arthritis-an inception cohort study. Arthritis Res Ther. 2021;23(1):169.
45. Van der Elst K, Verschueren P, De Cock D, De Groef A, Stouten V, Pazmino S, et al. One in five patients with rapidly and persistently controlled early rheumatoid arthritis report poor well-being after 1 year of treatment. RMD Open. 2020;6(1).
46. 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(8):2006-14.
47. Gwinnutt JM, Sharp CA, Symmons DPM, Lunt M, Verstappen SMM. Baseline patient reported outcomes are more consistent predictors of long-term functional disability than laboratory, imaging or joint count data in patients with early inflammatory arthritis: A systematic review. Semin Arthritis Rheum. 2018;48(3):384-98.
48. Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: Identifying reversible and irreversible components. Arthritis Rheum. 2006;54(9):2784-92.
49. Navarro-Compán V, Landewé R, Provan SA, Ødegård S, Uhlig T, Kvien TK, et al. Relationship between types of radiographic damage and disability in patients with rheumatoid arthritis in the EURIDISS cohort: a longitudinal study. Rheumatology (Oxford). 2015;54(1):83-90.
50. Radner H, Smolen JS, Aletaha D. Impact of comorbidity on physical function in patients with rheumatoid arthritis. Ann Rheum Dis. 2010;69(3):536-41.
51. McWilliams DF, Walsh DA. Pain mechanisms in rheumatoid arthritis. Clin Exp Rheumatol. 2017;35 Suppl 107(5):94-101.
52. Katz P. Causes and consequences of fatigue in rheumatoid arthritis. Current opinion in rheumatology. 2017;29(3):269-76.
53. Matcham F, Galloway J, Hotopf M, Roberts E, Scott IC, Steer S, et al. The impact of targeted Rheumatoid Arthritis pharmacological treatment on mental health: A systematic review and network meta-analysis. Arthritis Rheumatol. 2018.
54. Uhlig T, Haavardsholm EA, Kvien TK. Comparison of the Health Assessment Questionnaire (HAQ) and the modified HAQ (MHAQ) in patients with rheumatoid arthritis. Rheumatology (Oxford). 2006;45(4):454-8.

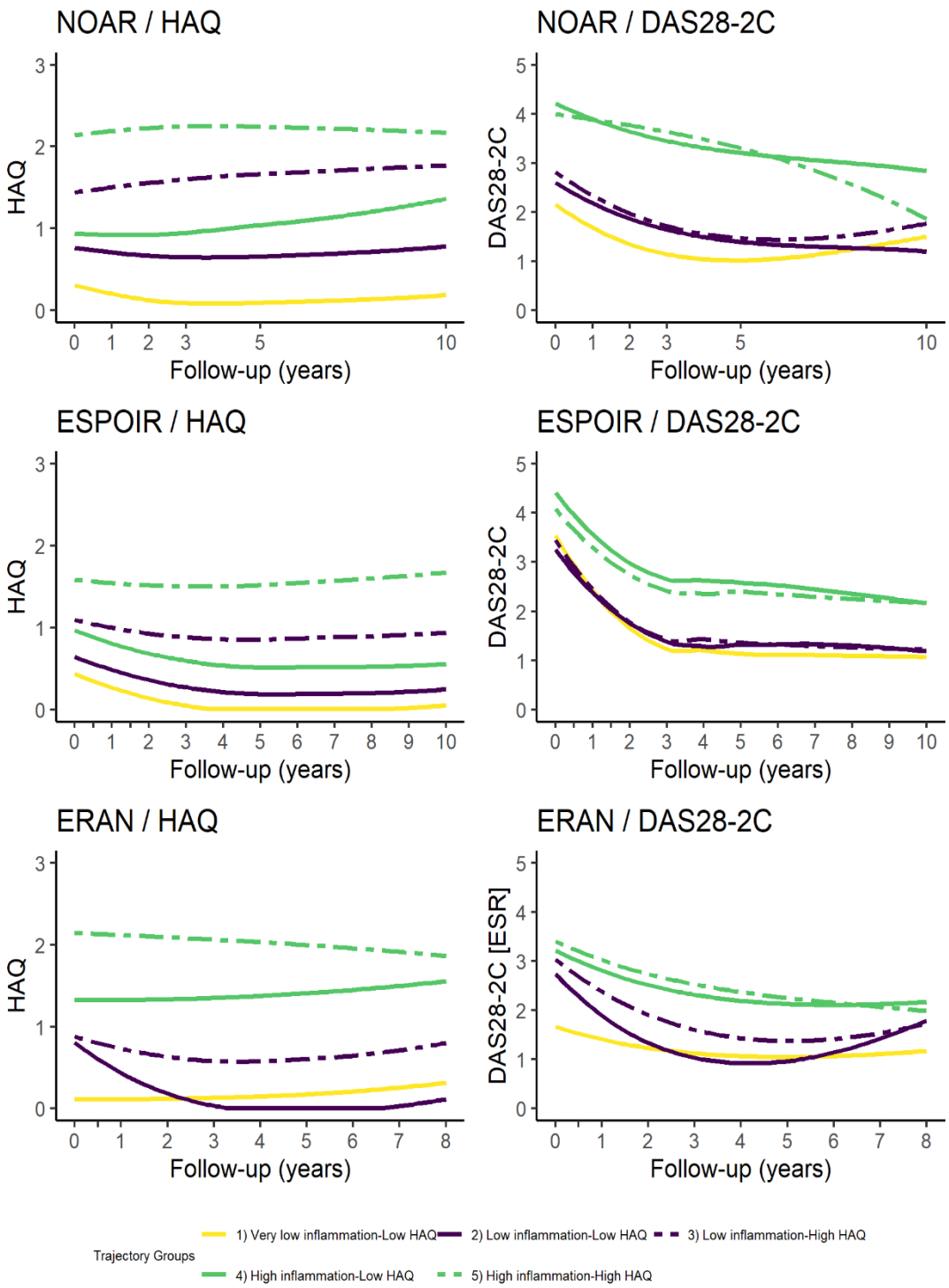


*Figure 1 – Trajectories of inflammation and disability over follow-up from the GBTM analysis*

*The HAQ component of each trajectory group's name is relative to the pair (i.e. in the "low inflammation pair" there is one group with low HAQ and one with high HAQ)*
*DAS28-2C = Two-component Disease Activity Score, ERAN = Early Rheumatoid Arthritis Network, ESPOIR = Étude et Suivi des Polyarthrites Indifférenciées Récentes, HAQ = Health Assessment Questionnaire, NOAR = Norfolk Arthritis Register*

*Table 1 – Baseline characteristics of the trajectory groups*

| | NOAR | | | | ESPOIR | | | | ERAN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Low inflammation | | High inflammation | | Low inflammation | | High inflammation | | Low inflammation | | High inflammation | |
| | Low HAQ | High HAQ | Low HAQ | High HAQ | Low HAQ | High HAQ | Low HAQ | High HAQ | Low HAQ | High HAQ | Low HAQ | High HAQ |
| N (%) | 295 (29.5%) | 199 (19.9%) | 104 (10.4%) | 115 (11.5%) | 228 (29.8%) | 127 (16.6%) | 136 (17.8%) | 89 (11.6%) | 87 (11.9%) | 208 (28.3%) | 211 (28.7%) | 119 (16.2%) |
| Age, years | 55.0 (14.5) | 62.8 (13.8)‡ | 56.6 (12.7) | 59.3 (15.2) | 46.6 (13.1) | 49.5 (9.9)† | 48.7 (12.8) | 52.0 (11.4)† | 54.1 (13.5) | 56.6 (14.5) | 57.5 (12.6) | 59.6 (13.5) |
| Women, N (%) | 192 (65.1%) | 138 (69.3%) | 79 (76.0%) | 86 (74.8%) | 173 (75.9%) | 109 (85.8%)† | 99 (72.8%) | 79 (88.8%)† | 45 (51.7%) | 155 (74.5%)‡ | 151 (71.6%) | 102 (85.7%)† |
| Symptom duration, months | 7.8 (5.0) | 7.8 (5.4) | 8.7 (5.5) | 8.7 (5.8) | 3.4 (1.5) | 3.7 (2.1) | 3.3 (1.7) | 3.7 (1.8) | 10.0 (4.5) | 10.0 (5.1) | 10.2 (5.2) | 10.1 (5.7) |
| BMI | 26.9 (5.0) | 27.7 (5.7) | 28.0 (5.7) | 29.3 (6.2) | 24.9 (4.5) | 25.7 (5.3) | 25.7 (4.6) | 26.5 (4.7) | 26.6 (5.1) | 27.3 (5.6) | 28.6 (5.2) | 29.2 (6.9) |
| BMI categories § | | | | | | | | | | | | |
| Underweight | 2 (0.7%) | 6 (3.0%) | 1 (1.0%) | 3 (2.6%) | 8 (3.5%) | 3 (2.4%) | 2 (1.5%) | 3 (3.4%) | 1 (1.1%) | 4 (1.9%) | 0 (0%) | 3 (2.5%) |
| Normal weight | 114 (38.6%) | 60 (30.2%) | 33 (31.7%) | 26 (22.6%) | 133 (58.3%) | 63 (49.6%) | 66 (48.5%) | 37 (41.6%) | 31 (35.6%) | 69 (33.2%) | 43 (20.4%) | 25 (21.0%) |
| Overweight | 107 (36.3%) | 70 (35.2%) | 36 (34.6%) | 36 (31.3%) | 56 (24.6%) | 34 (26.8%) | 45 (33.1%) | 28 (31.5%) | 30 (34.5%) | 69 (33.2%) | 83 (39.3% | 42 (35.3%) |
| Obese | 70 (23.7%) | 61 (30.7%) | 32 (30.8%) | 46 (40.0%) | 31 (13.6%) | 25 (19.7%) | 23 (16.9%) | 21 (23.6%) | 15 (17.2%) | 51 (24.5%) | 62 (29.4%) | 43 (36.1%) |
| Missing | 2 (0.7%) | 2 (1.0%) | 2 (1.9%) | 4 (3.5%) | 0 (0%) | 2 (1.6%) | 0 (0%) | 0 (0%) | 10 (11.5%) | 15 (7.2%) | 23 (10.9%) | 6 (5.0%) |
| Smoking, N (%) | | | | | | | | | | | | |
| Smoker | 68 (23.1%) | 41 (20.6%) | 23 (22.1%) | 31 (27.0%) | 110 (48.2%) | 62 (48.8%) | 65 (47.8%) | 42 (47.2%) | 28 (32.2%) | 62 (29.8%) | 73 (34.6%) | 43 (36.1%) |
| Non-smoker | 189 (64.1%) | 137 (68.8%) | 70 (67.3%) | 77 (67.0%) | 118 (51.8%) | 65 (51.2%) | 71 (52.2%) | 47 (52.8%) | 59 (67.8%) | 143 (68.8%) | 136 (64.5%) | 76 (63.9%) |
| Missing | 38 (12.9%) | 21 (10.6%) | 11 (10.6%) | 7 (6.1%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 3 (1.4%) | 2 (1.0%) | 0 (0%) |
| DAS28-CRP (ERAN:ESR) | 3.6 (1.1) | 3.8 (1.2) | 4.5 (1.1) | 4.9 (1.3)† | 4.1 (1.1) | 4.5 (1.0)† | 5.1 (1.1) | 5.1 (1.1) | 4.4 (1.1) | 4.6 (1.2) | 4.9 (1.4) | 5.5 (1.2)‡ |
| DAS28-2C (ERAN: ESR) | 2.8 (1.4) | 3.0 (1.5) | 4.4 (1.3) | 4.1 (1.6) | 3.6 (1.2) | 3.8 (1.1) | 4.7 (1.3) | 4.5 (1.4) | 3.1 (1.2) | 3.3 (1.2) | 3.4 (1.4) | 3.6 (1.3) |
| HAQ | 0.8 (0.6) | 1.5 (0.5)‡ | 1.0 (0.5) | 2.1 (0.5)‡ | 0.8 (0.6) | 1.2 (0.6)‡ | 1.1 (0.6) | 1.7 (0.6)‡ | 1.0 (0.5) | 0.9 (0.6) | 1.3 (0.5) | 2.1 (0.4)‡ |
| Pain VAS (ERAN: SF36-P) | 36.4 (24.8) | 46.8 (25.6)‡ | 41.2 (23.8) | 69.2 (22.2)‡ | 33.5 (25.7) | 45.0 (28.5)‡ | 41.1 (28.0) | 50.2 (26.1)† | 47.8 (23.0) | 47.0 (22.4) | 40.9 (21.7) | 24.3 (18.9)‡ |
| Fatigue VAS (ERAN: SF36-V) | 42.2 (27.3) | 53.0 (25.7)‡ | 47.8 (28.6) | 71.1 (21.9)‡ | 42.9 (26.6) | 53.5 (28.6)‡ | 51.2 (26.0) | 66.2 (23.3)‡ | 46.6 (22.8) | 45.0 (18.9) | 36.9 (18.7) | 27.4 (19.0)‡ |
| AIMS Depression (ERAN: SF36-MH) | 2.89 (1.91) | 3.54 (1.97)† | 3.01 (1.72) | 4.85 (2.16)‡ | 3.07 (1.92) | 3.96 (2.01)‡ | 4.03 (1.94) | 5.21 (2.35)‡ | 69.6 (19.6) | 69.0 (18.0) | 63.0 (18.0) | 54.1 (20.7)‡ |
| AIMS Anxiety | 4.00 (2.03) | 4.48 (2.02)† | 4.23 (1.68) | 5.57 (2.17)‡ | 4.47 (2.32) | 5.32 (2.14)‡ | 5.07 (2.32) | 6.11 (2.30)† | - | - | - | - |
| RF, N (%)) | | | | | | | | | | | | |
| Positive | 121 (41.0%) | 86 (43.2%) | 47 (45.2%) | 39 (33.9%) | 100 (43.9%) | 56 (44.1%) | 74 (54.4%) | 40 (44.9%) | 41 (47.1%) | 107 (51.4%) | 110 (52.1%) | 77 (64.7%)† |
| Negative | 164 (55.6%) | 108 (54.3%) | 53 (51.0%) | 68 (59.1%) | 128 (56.1%) | 71 (55.9%) | 62 (45.6%) | 49 (55.1%) | 34 (39.1%) | 66 (31.7%) | 86 (40.8%) | 27 (22.7%) |
| Missing | 10 (3.4%) | 5 (2.5%) | 4 (3.9%) | 8 (7.0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 12 (13.8%) | 35 (16.8%) | 15 (7.1%) | 15 (12.6%) |
| Anti-CCP, N (%)) | | | | | | | | | | | | |
| Positive | 90 (30.5%) | 73 (36.7%) | 44 (42.3%) | 35 (30.4%) | 84 (36.8%) | 42 (33.1%) | 73 (53.7%) | 36 (40.4%) | - | - | - | - |
| Negative | 187 (63.4%) | 111 (55.8%) | 47 (45.2%) | 66 (57.4%) | 144 (63.2%) | 85 (66.9%) | 63 (46.3%) | 53 (59.6%) | | | | |
| Missing | 18 (6.1%) | 15 (7.5%) | 13 (12.5%) | 14 (12.2%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | | | | |
| csDMARD, N (%) | | | | | | | | | | | | |
| Yes | 155 (52.5%) | 107 (53.8%) | 57 (54.8%) | 61 (53.0%) | 13 (5.7%) | 5 (3.9%) | 11 (8.1%) | 3 (3.4%) | 57 (65.5%) | 124 (59.6%) | 131 (62.1%) | 76 (63.9%) |
| No | 140 (47.5%) | 92 (46.2%) | 47 (45.2%) | 54 (47.0%) | 215 (94.3%) | 122 (96.1%) | 125 (91.9%) | 86 (96.6%) | 30 (34.5%) | 84 (40.4%) | 80 (37.9%) | 43 (36.1%) |
| Comorbidities [ERAN: RDCI) | | | | | | | | | | | | |
| 0 | 80 (27.1%) | 34 (17.1%) | 23 (22.1%) | 14 (12.2%)† | 130 (57.0%) | 70 (55.1%) | 79 (58.1%) | 39 (43.8%) | 0.41 (0.72) | 0.64 (0.92)† | 0.79 (1.1) | 0.93 (1.0) |
| 1 | 77 (26.1%) | 46 (23.1%) | 21 (20.2%) | 18 (15.7%) | 61 (26.8%) | 36 (28.3%) | 30 (22.1%) | 24 (27.0%) | | | | |
| ≥2 | 57 (19.3%) | 48 (24.1%) | 14 (13.5%) | 32 (27.8%) | 37 (16.2%) | 21 (16.5%) | 27 (19.9%) | 26 (29.2%) | | | | |
| Missing | 81 (27.5%) | 71 (35.7%) | 46 (44.2%) | 51 (44.3%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | | | | |
| Erosions | | | | | | | | | | | | |
| Yes | - | - | - | - | 20 (8.8%) | 20 (15.7%) | 25 (18.4%) | 17 (19.1%) | 19 (21.8%) | 60 (28.8%) | 52 (24.6%) | 30 (25.2%) |
| No | | | | | 208 (91.2%) | 107 (84.3%) | 111 (81.6%) | 72 (80.9%) | 66 (75.9%) | 139 (66.8%) | 149 (70.6%) | 84 (70.6%) |
| Missing | | | | | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 2 (2.3%) | 9 (4.3%) | 10 (4.7%) | 5 (4.2%) |

*† p<0.05, ‡ p<0.001 - comparing high HAQ vs low HAQ groups within each inflammation pair; t-tests used to compare continuous variables, chi² for categorical variables; § BMI categories: Underweight (BMI <18.5), Normal weight (BMI ≥18.5 & <25), Overweight (BMI ≥25 & <30), Obese (BMI ≥30); * depression and anxiety measured with the Arthritis Impact Measurement Scales for NOAR and ESPOIR, SF36 mental health for ERAN*

*See supplementary table 7 for proportions of missing data for each variable*

*AIMS = Arthritis Impact Measurement Scales, Anti-CCP = anti-cyclic citrullianted protein antibodies, BMI = body mass index, csDMARD = conventional synthetic disease modifying anti-rheumatic drugs, DAS28 = Disease Activity Score 28, DAS28-2C – Disease Activity Score 28 – 2 components, ESPOIR = Étude et Suivi des Polyarthrites Indifférenciées Récentes, ESR = erythrocyte sedimentation rate, ERAN = Early Rheumatoid Arthritis Network, HAQ = Health Assessment Questionnaire, N = Number, NOAR = Norfolk Arthritis Register, RF = rheumatoid factor, SD = standard deviation, SF36 = Short-Form 36 (SF36-MH = SF36-mental healthSF36-P = pain scale, SF36-V = vitality scale), VAS = visual analogue scale*

*Table 2 – Baseline predictors of high HAQ group membership compared with corresponding low HAQ group in each inflammation pair*

| | Low inflammation | | | High inflammation | | |
|---|---|---|---|---|---|---|
| **Variable** | **NOAR, OR (95% CI)** | **ESPOIR, OR (95% CI)** | **ERAN, OR (95% CI)** | **NOAR, OR (95% CI)** | **ESPOIR, OR (95% CI)** | **ERAN, OR (95% CI)** |
| Age, years | 1.07 (1.05, 1.09) | 1.03 (1.01, 1.06) | 1.02 (0.99, 1.04) | 1.05 (1.02, 1.09) | 1.03 (1.00, 1.06) | 1.03 (1.01, 1.06) |
| Female vs male | 1.59 (1.00, 2.52) | 2.34 (1.19, 4.60) | 4.15 (2.25, 7.68) | 0.62 (0.25, 1.58) | 3.48 (1.42, 8.57) | 3.61 (1.80, 7.25) |
| Symptom duration, months | 1.01 (0.97, 1.05) | 1.09 (0.94, 1.26) | 0.99 (0.93, 1.05) | 0.97 (0.90, 1.04) | 1.24 (1.03, 1.49) | 1.01 (0.96, 1.06) |
| Current smoker vs non-smoker | 1.14 (0.63, 2.06) | 1.25 (0.76, 2.06) | 1.01 (0.56, 1.83) | 1.86 (0.65, 5.28) | 1.40 (0.72, 2.72) | 1.20 (0.67, 2.16) |
| BMI | 1.03 (0.99, 1.07) | 1.02 (0.97, 1.08) | 1.02 (0.96, 1.09) | 1.06 (0.99, 1.13) | 1.05 (0.98, 1.13) | 1.02 (0.98, 1.07) |
| Pain (VAS†: NOAR / ESPOIR; SF36‡: ERAN) | 1.20 (1.09, 1.31) | 1.12 (1.02, 1.23) | 0.88 (0.62, 1.26) | 1.48 (1.23, 1.78) | 1.06 (0.93, 1.20) | 0.34 (0.22, 0.53) |
| Fatigue (VAS†: NOAR / ESPOIR; SF36‡: ERAN) | 1.15 (1.02, 1.28) | 1.06 (0.96, 1.17) | 0.90 (0.58, 1.38) | 1.22 (0.96, 1.54) | 1.22 (1.06, 1.41) | 0.79 (0.54, 1.14) |
| Depression (AIMS2: NOAR / ESPOIR; SF36‡: ERAN) | 1.11 (0.91, 1.34) | 1.24 (1.07, 1.44) | 1.15 (0.74, 1.80) | 1.40 (1.01, 1.94) | 1.14 (0.95, 1.37) | 0.87 (0.62, 1.23) |
| Anxiety (AIMS2: NOAR / ESPOIR) | 1.04 (0.86, 1.26) | 1.01 (0.88, 1.15) | - | 0.99 (0.67, 1.47) | 1.08 (0.92, 1.27) | - |
| RF | 0.85 (0.52, 1.41) | 1.22 (0.65, 2.30) | 1.53 (0.85, 2.75) | 0.55 (0.21, 1.44) | 0.96 (0.40, 2.31) | 2.01 (1.07, 3.78) |
| Anti-CCP | 1.79 (1.03, 3.11) | 0.84 (0.43, 1.64) | - | 1.21 (0.46, 3.21) | 0.86 (0.35, 2.11) | - |
| Taking csDMARDs | 0.85 (0.56, 1.31) | 0.66 (0.20, 2.15) | 0.77 (0.43, 1.36) | 0.71 (0.30, 1.69) | 0.60 (0.13, 2.77) | 0.98 (0.55, 1.74) |
| Comorbidities (RDCI: ERAN) | | | | | | |
| 1 vs 0 comorbidities | 1.18 (0.62, 2.26) | 0.96 (0.54, 1.69) | 1.42 (0.97, 2.08) | 0.93 (0.31, 2.76) | 1.22 (0.54, 2.75) | 1.12 (0.86, 1.46) |
| 2 vs 0 comorbidities | 1.31 (0.73, 2.35) | 0.93 (0.45, 1.91) | - | 1.66 (0.53, 5.22) | 1.18 (0.51, 2.74) | - |
| Erosions vs no erosions | - | 1.75 (0.83, 3.68) | 1.68 (0.88, 3.22) | - | 1.00 (0.44, 2.29) | 1.00 (0.52, 1.92) |

*† VAS measured in centimetres, ‡ SF36 – higher scores indicate better status. SF36 scores were standardised, therefore odds ratio represents a standard deviation increase in SF36*

*AIMS2 = Arthritis Impact Measurement Scales 2, Anti-CCP = Anti-cyclic citrullinated peptide antibody, BMI = body mass index, CI = confidence interval, csDAMRD = conventional synthetic disease modifying anti-rheumatic drug, ERAN = Early Rheumatoid Arthritis Network, ESPOIR = Étude et Suivi des Polyarthrites Indifférenciées Récentes, HAQ = Health Assessment Questionnaire, NOAR = Norfolk Arthritis Register, OR = odds ratio, RDCI = Rheumatic Disease Comorbidity Index, RF = rheumatoid factor, SF36 = Short form 36, VAS = visual analogue scale*

*Table 3 – Comparison of the outcomes over ten years between the high and low HAQ trajectories, stratified by inflammation pair and cohort, mean difference (95% CI)†*

| | NOAR | | ESPOIR | | ERAN | |
|---|---|---|---|---|---|---|
| | **Low-inflammation: high HAQ vs low HAQ** | **High-inflammation: high HAQ vs low HAQ** | **Low-inflammation: high HAQ vs low HAQ** | **High-inflammation: high HAQ vs low HAQ** | **Low-inflammation: high HAQ vs low HAQ** | **High-inflammation: high HAQ vs low HAQ** |
| *Trajectory components* | | | | | | |
| DAS28-2C [NOAR/ESPOIR: CRP, ERAN: ESR] | 0.17 (0.03, 0.30) | -0.13 (-0.35, 0.10) | 0.04 (-0.07, 0.15) | -0.17 (-0.36, 0.02) | 0.40 (0.22, 0.58) | 0.11 (-0.10, 0.32) |
| HAQ | 0.86 (0.81, 0.91) | 1.22 (1.13, 1.30) | 0.63 (0.59, 0.66) | 0.94 (0.87, 1.01) | 0.39 (0.33, 0.44) | 0.72 (0.66, 0.77) |
| *Additional PROMs* | | | | | | |
| Pain VAS / SF36-pain § (ERAN) | 18.0 (15.1, 21.0) | 27.0 (23.1, 31.0) | 11.0 (8.3, 13.8) | 21.3 (17.0, 25.5) | -13.7 (-17.2, -10.2) | -15.8 (-19.0, -12.5) |
| Fatigue VAS / SF36-vitality § [ERAN] | 19.5 (15.9, 23.0) | 23.4 (18.1, 28.6) | 14.1 (10.6, 17.7) | 22.1 (17.3, 26.9) | -10.3 (-14.4, -6.2) | -10.6 (-14.0, -7.3) |
| AIMS Depression / SF36-mental health § [ERAN] | 1.0 (0.7, 1.3) | 1.9 (1.5, 2.4) | 1.2 (0.9, 1.5) | 1.8 (1.3, 2.2) | -7.0 (-10.6, -3.4) | -11.8 (-15.3, -8.2) |
| AIMS Anxiety | 0.8 (0.5, 1.1) | 1.7 (1.2, 2.1) | 1.2 (0.8, 1.6) | 1.7 (1.2, 2.1) | - | - |
| *DAS28 & components* | | | | | | |
| DAS28-CRP / DAS28-ESR [ERAN] | 0.50 (0.37, 0.64) | 0.69 (0.48, 0.91) | 0.29 (0.18, 0.40) | 0.44 (0.25, 0.63) | 0.48 (0.28, 0.69) | 0.39 (0.17, 0.61) |
| Swollen joint count (28) | 0.35 (0.05, 0.66) | 0.72 (-0.25, 1.70) | 0.06 (-0.16, 0.29) | -0.57 (-1.12, -0.01) | 1.13 (0.56, 1.69) | 0.41 (-0.37, 1.18) |
| Tender joint count (28) | 2.86 (2.06, 3.67) | 7.22 (5.56, 8.88) | 1.61 (1.03, 2.19) | 3.41 (2.15, 4.67) | 1.61 (0.81, 2.41) | 3.14 (1.92, 4.37) |
| CRP / ESR [ERAN] | 1.52 (-0.09, 3.12) | -4.10(-10.11, 1.92) | 0.47 (-0.68, 1.61) | 2.94 (0.09, 5.78) | 2.14 (-1.11, 5.40) | 0.57 (-3.06, 4.19) |
| *Comorbidities* | | | | | | |
| Presence of comorbidities‡ [ERAN: RDCI] | OR 1.17 (0.80, 1.71) | OR 1.61 (0.93, 2.80) | OR 1.41 (0.82, 2.42) | OR 1.91 (0.87, 4.21) | 0.28 (0.07, 0.49)¥ | 0.07 (-0.17, 0.31)¥ |

*† adjusted for age and gender, § SF36 = higher scores indicate lower pain / fatigue / better mental health, ‡ NOAR and ESPOIR analysed using generalised estimating equations analysis (GEE),* ¥ mean difference rather than OR for comorbidities data in ERAN, as the RDCI, which is a continuous measure, was used.

*CI = confidence interval DAS28 = Disease Activity Score, DAS28-2C = 2 component DAS28, ESPOIR = Étude et Suivi des Polyarthrites Indifférenciées Récentes, ERAN = Early Rheumatoid Arthritis Network, HAQ = Health Assessment Questionnaire, NOAR = Norfolk Arthritis Register, OR = odds ratio, PROMs = Patient Reported Outcome Measures, RDCI = Rheumatic Disease Comorbidity Index, VAS = Visual Analogue Scale*

*Table 4 – Interactions between high / low HAQ trajectory group and patient reported outcomes predicting HAQ score, stratified by inflammation pair and cohort*

| | Pain | | | Fatigue | | | Anxiety | | | Depression | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **VAS Pain§ coef.** | **High/low HAQ coef.** | **Interaction** | **VAS Fatigue§ coef.** | **High/low HAQ coef.** | **Interaction** | **Anxiety coef.** | **High/low HAQ coef.** | **Interaction** | **Depression coef.** | **High/low HAQ coef.** | **Interaction** |
| **NOAR** | | | | | | | | | | | | |
| Low inflammation | 0.09 | 0.62 | -0.02 | 0.05 | 0.55 | 0.00 | 0.05 | 0.58 | 0.00 | 0.06 | 0.57 | 0.002 |
| pair | (0.08, 0.11) | (0.51, 0.72) | (-0.04, -0.01) | (0.04, 0.06) | (0.44, 0.66) | (-0.02, 0.01) | (0.03, 0.07) | (0.48, 0.68) | (-0.02, 0.02) | (0.04, 0.07) | (0.48, 0.66) | (-0.02, 0.02) |
| High inflammation | 0.08 | 0.84 | -0.03 | 0.06 | 0.89 | -0.02 | 0.05 | 0.85 | -0.02 | 0.05 | 0.88 | -0.03 |
| pair | (0.06, 0.10) | (0.68, 0.99) | (-0.05, -0.002) | (0.04, 0.08) | (0.71, 1.07) | (-0.05, 0.003) | (0.01, 0.08) | (0.69, 1.02) | (-0.05, 0.01) | (0.01, 0.09) | (0.70, 1.07) | (-0.06, 0.01) |
| **ESPOIR** | | | | | | | | | | | | |
| Low inflammation | 0.06 | 0.39 | -0.0004 | 0.03 | 0.35 | 0.01 | 0.03 | 0.32 | 0.02 | 0.07 | 0.31 | 0.02 |
| pair | (0.05, 0.07) | (0.34, 0.44) | (-0.01, 0.01) | (0.03, 0.04) | (0.30, 0.41) | (0.001, 0.02) | (0.03, 0.04) | (0.26, 0.38) | (0.01, 0.03) | (0.06, 0.08) | (0.25, 0.37) | (0.01, 0.04) |
| High inflammation | 0.09 | 0.69 | -0.04 | 0.07 | 0.66 | -0.02 | 0.07 | 0.69 | -0.02 | 0.11 | 0.69 | -0.03 |
| pair | (0.08, 0.10) | (0.60, 0.78) | (-0.06, -0.03) | (0.06, 0.08) | (0.56, 0.77) | (-0.03, -0.01) | (0.06, 0.08) | (0.58, 0.80) | (-0.04, -0.003) | (0.09, 0.12) | (0.59, 0.79) | (-0.05, -0.01) |
| **ERAN** | **SF36 - Pain** | | | **SF36 - Vitality** | | | **SF36 – Mental Health** | | | | | |
| Low inflammation | -0.01 | 0.12 | 0.00 | -0.008 | 0.23 | 0.0003 | -0.01 | 0.23 | 0.001 | - | - | - |
| pair | (-0.01, -0.01) | (-0.04, 0.27) | (-0.001, 0.003) | (-0.01, -0.005) | (0.07, 0.38) | (-0.002, 0.003 | (-0.01, -0.004) | (-0.01, 0.47) | (-0.002, 0.004) | | | |
| High inflammation | -0.01 | 0.41 | 0.0001 | -0.007 | 0.51 | -0.001 | -0.004 | 0.57 | -0.002 | - | - | - |
| pair | (-0.01, -0.01) | (0.32, 0.51) | (-0.002, 0.003) | (-0.009, -0.004) | (0.40, 0.62) | (-0.004, 0.002) | (-0.01, -0.001) | (0.37, 0.77) | (-0.004, 0.001) | | | |

*Analyses controlling for age, gender, baseline comorbidity and BMI, and lagged HAQ, missing data imputed using multiple imputation*

*§ VAS in cm*

*AIMS = Arthritis Impact Measurement Scales, ERAN = Early rheumatoid arthritis network, ESPOIR = Étude et Suivi des Polyarthrites Indifférenciées Récentes, HAQ = Health Assessment Questionnaire, NOAR = Norfolk Arthritis Register, SF36 = short form (36), VAS = visual analogue scale*

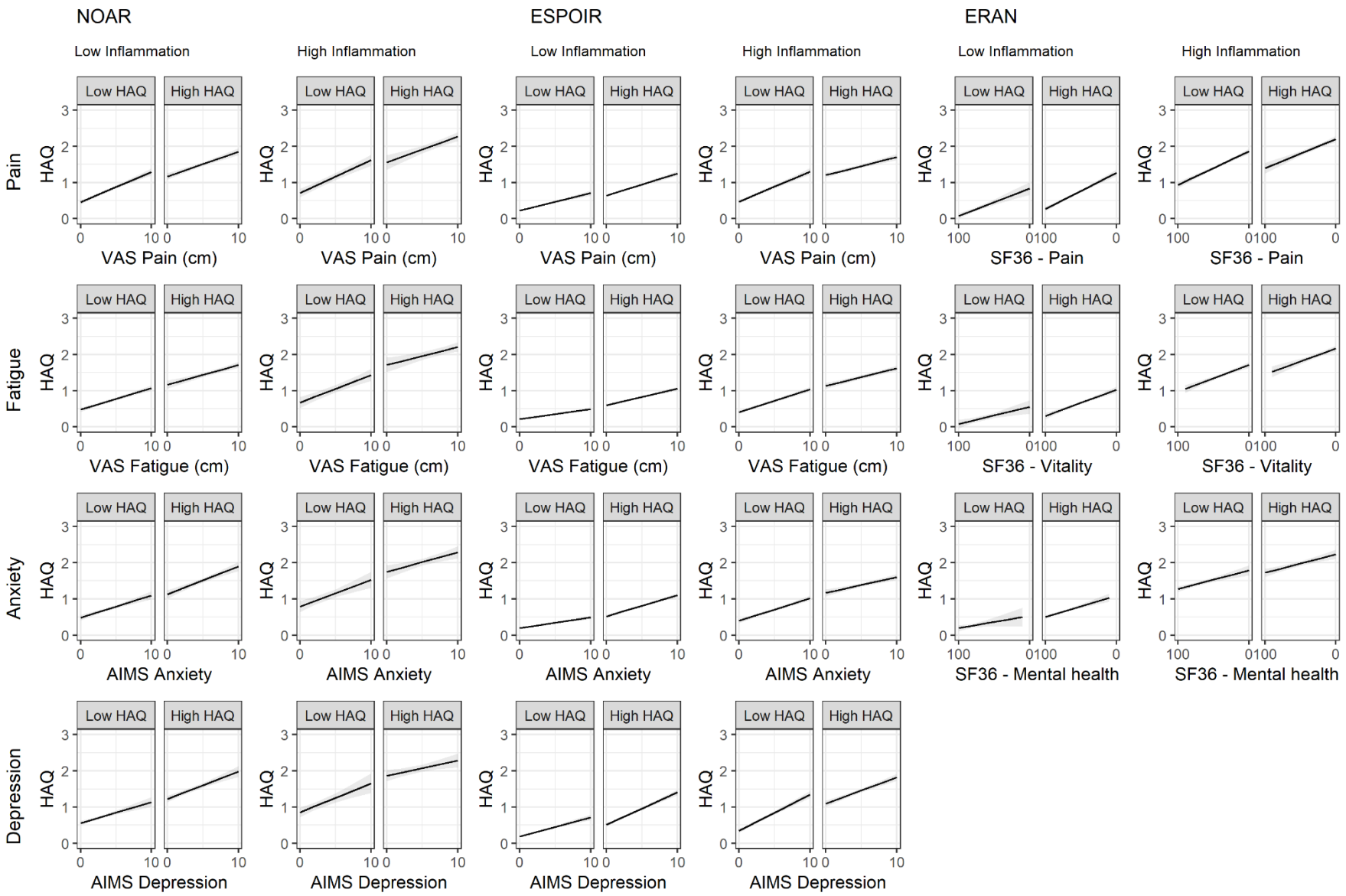


*Figure 2 – Interactions between high and low HAQ group status and patient report outcomes predicting HAQ score, stratified by inflammation status and cohort (unimputed results)*

*Figure indicates that, for many PROMS, the association between PROM score and HAQ score was stronger in the lower-HAQ trajectory within each inflammation pair (i.e. the slope is steeper in the low-HAQ trajectory compared with the high-HAQ trajectory).*

*X axis of the ERAN analyses reversed, as higher scores on the SF36 indicate better outcomes*

*AIMS = Arthritis Impact Measurement Scales, ERAN = Early rheumatoid arthritis network, ESPOIR = Étude et Suivi des Polyarthrites Indifférenciées Récentes, HAQ = Health Assessment Questionnaire, NOAR = Norfolk Arthritis Register, SF36 = short form (36), VAS = visual analogue scale*