The role of comorbidities alongside patient and disease characteristics on long-term disease activity in RA using UK inception cohort data

Amanda D Busby1, James Wason2, Arthur G Pratt3,4, Adam Young1, John D Isaacs3,4, Elena Nikiphorou5,6

1Centre for Health Services and Clinical Research, Life and Medical Sciences, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK
2Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK
3Newcastle University Translational and Clinical Research Institute, Faculty of Medical Sciences, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK
4Musculoskeletal Unit, Newcastle upon Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Freeman Road, Newcastle upon Tyne, NE7 7DN, UK
5Centre for Rheumatic Diseases, King’s College London, London, UK
6Rheumatology Department, King’s College Hospital, London, UK

Corresponding author
Amanda D Busby
a.busby@herts.ac.uk
ORCiD ID: https://orcid.org/0000-0002-0545-0276

Word Count: 3419

Keywords:
Sociodemographic factors, Comorbidity, Multimorbidity, Disease Activity, Rheumatoid Arthritis.

Key Messages:
1. Sociodemographic factors are associated with worse disease activity in RA, independently of clinical measures at first presentation.
2. Identifying patients at increased risk of moderate/high disease activity at initial review enables targeted management.
3. Unlike sociodemographic factors and other clinical measures at baseline, comorbidities were not associated with worse disease activity.
Abstract

Objectives
Control of disease activity in Rheumatoid Arthritis (RA) is a crucial part of its management to prevent long-term joint damage and disability. This study aimed to identify early predictors of poor disease activity at 5 and 10 years, focusing on comorbidities and clinical/sociodemographic factors at first presentation.

Methods
Patients from two UK-based RA cohorts were classified into two groups; low (<3.2) and moderate/high (≥3.2) Disease Activity Score (DAS28) at five/10 years. Clinical variables (e.g., rheumatoid nodules, erosions), sociodemographic factors (e.g., ethnicity, deprivation) and comorbidities were recorded at baseline and yearly thereafter. The Rheumatic Diseases Comorbidity Index (RDCI) quantified patient comorbidity burden. Binary logistic regression models (outcome low versus moderate/high DAS28) were fitted using multiple imputation.

Results
2,701 patients living with RA were recruited (mean age 56.1 years, 66.9% female); five-year data were available for 1,718 (63.4%) patients and 10-year for 820 (30.4%). Baseline RDCI was not associated with DAS28 at five (OR 1.05, 95% CI 0.91 to 1.22) or 10 years (OR 0.99, 95% CI 0.75 to 1.31) in multivariable analyses. Sociodemographic factors (female gender, worse deprivation) and poorer baseline HAQ-DI were associated with DAS28≥3.2 at both timepoints. Being seropositive was associated with five-year DAS28≥3.2.

Conclusion
This study demonstrates an association between sociodemographic and clinical factors and long-term RA disease activity, in models adjusting for comorbidity burden. The findings call for more holistic and targeted patient management in patients with RA and provide insights for more individualised management plans even on first presentation to rheumatology.
Introduction
Managing disease activity is a vital aspect of care for patients living with rheumatoid arthritis (RA). Active disease, measured using patient-reported outcomes (PROs), clinical assessment or a combination of the two by use of a composite score(1), can lead to long-term joint damage(2), causing permanent discomfort and/or disability resulting in orthopaedic surgery(3). Because of this, the European Alliance of Associations for Rheumatology (EULAR)(4) guidelines recommend focus on disease activity for measuring RA patients’ health outcomes. In particular, one of EULAR’s four overarching principles targets clinical remission, defined as the absence of significant inflammatory disease activity, or low disease activity as an acceptable alternative(5).

In EULAR’s most recent guidelines (updated in 2014), attention was drawn to the role of comorbidities in RA disease management noting that they had “potential to preclude treatment intensification”(5). Aside from this, recent work by the EULAR Task Force on difficult-to-treat RA recognises that comorbidities can influence the assessment of inflammatory activity in patients with RA(6). In fact, two points to consider are highlighted through this work and are worth addressing here. The first calls for caution when using composite indices and clinical evaluation in the presence of comorbidities, in particular obesity and fibromyalgia, since these may directly heighten inflammatory activity and/or overestimate disease activity respectively(7–10). It should be noted that comorbidities (such as infections and malignancies) can also contribute to the difficult-to-treat state in RA(11,12), through misclassification of inflammatory RA activity(13). Another point of consideration highlights that comorbidities that impact quality of life, either independently or by limiting RA treatment options, should be carefully considered and managed.

Most of the evidence to date is in RA more generally, suggesting a role of comorbidities on physical disability(14), although it remains unclear what effect they have on actual disease activity. As for difficult-to-treat disease, obesity is once again implicated in higher disease activity(15,16), but the role of other comorbidities on disease activity per se is generally less well known. These observations, together with a knowledge that comorbidities are more prevalent in RA patients than the general population(17), and prominent even on first disease presentation(18,19), have formed the rationale for this work.

The aim of this study was to identify early predictors of poor disease activity in RA at five and 10 years, focusing on comorbidities and exploring associations with other clinical and sociodemographic factors.
Methods

Study Population
Data collected as part of two consecutive prospective cohorts were used in this analysis; the Early Rheumatoid Arthritis Study (ERAS, 1986-2001) and Early Rheumatoid Arthritis Network (ERAN, 2002-2012). These cohorts were designed to be similar so can be analysed together and have previously been described in detail(20). Recruitment for ERAS took place across nine hospitals in England and for ERAN through 23 centres in England, Wales and the Republic of Ireland. ERAS patients gave informed consent as required at time of enrolment (prior to Good Clinical Practice (GCP) implementation), approved by the East Hertfordshire Local Research Ethics Committee. ERAN patients gave informed, written consent, approved by the Trent Research Ethics Committee. Patients were followed up 3-6 months after baseline visit, yearly thereafter. Dates and causes of deaths for patients were obtained from NHS Digital.

Baseline Measures

Comorbidities
The primary focus of this study was whether baseline comorbidity, measured using the Rheumatic Disease Comorbidity Index (RDCI)(21) predicted future disease activity. The RDCI, used as a continuous variable, was developed as a rheumatic-specific measure of comorbidity, a predictor both of death and of functional outcomes. RDCI scores range from 0 to 9, with higher scores indicating a larger comorbidity burden. Eight disease categories are included, consisting of 11 illness types: lung disease, heart attack/other cardiovascular/stroke, fracture, depression, cancer, ulcer/stomach disease, hypertension, diabetes. The category definitions use the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes. The overall RDCI score or individual disease category were included in all analyses a priori.

Sociodemographic and Clinical Variables
Secondary variables included standard demographic data (age at onset (to the nearest year), gender, ethnicity) were collected at baseline. Patient postcodes, recorded at baseline, were used to derive the Index of Multiple Deprivation (IMD)(22). Quintiles were calculated for each IMD measure and patients assigned a group from 1 (most deprived) to 5 (least deprived), such that comparisons could be made across all nations. Clinical data (body mass index (BMI), presence of erosions, presence of rheumatoid nodules, haemoglobin levels, presence of rheumatoid factor, presence of anti-cyclic citrullinated peptide autoantibodies (anti-CCP) and erythrocyte sedimentation rate (ESR)) were collected at baseline and then at each subsequent appointment. Patients were asked to complete
the health assessment questionnaire disability index (HAQ-DI) at each visit to ascertain their own views of functional disability. Time between onset of RA symptoms to first rheumatology appointment was recorded, measured in months. Time from onset to the commencement of disease modifying anti-rheumatic drugs (DMARD) for those prescribed was also recorded, again in months.

Some variables were reclassified to aid interpretation, for example smoking status (ever/never), serostatus (positive/negative, where positive included rheumatoid factor and/or anti-CCP), ethnicity (white/minority ethnic groups).

**Outcomes**

**Five-Year and 10-Year Disease Activity**

During data collection for ERAS, disease activity was measured using the Disease Activity Score (DAS)(23). As per standard practice at the time, this comprised four components, a combination of patient (subjective) and clinician (objective) measures. These were: the Ritchie articular index (grading tenderness of 52 joints from 0-3), a general health assessment on a visual analogue scale (from 0, good health to 100, poor health), the 44 swollen joint count and ESR. These were combined to give DAS values which were rounded to scores ranging from 1 (low disease activity) to 9 (high disease activity).

The DAS28 was subsequently developed(24) and was used to measure disease activity in ERAN patients, with simplified joint scores including 28 joints rather than 44/52. Alternative formulas enabled DAS28 to be calculated if no measure of general health was available or if C-reactive protein (CRP) was recorded in place of ESR. DAS28 scores range from 2 to 10, so are not directly interchangeable with DAS scores. ERAS DAS scores were therefore transformed using a tested formula(25) such that they could be compared directly to ERAN DAS28 values, as used in other studies. This formula, which uses all separate components of the DAS, estimates the scores with a higher level of precision than the transformation formula used previously, which tended to overestimate the true DAS28 score, particularly at the higher end of the scale.

DAS28 scores were dichotomised into remission/low (<3.2) or moderate/high disease activity (≥3.2) (26), a well-defined and clinically relevant cut-point. The primary outcome for this study was five-year DAS28≥3.2, with DAS28 at 10 years studied as secondary analysis.
Statistical Analysis

The statistical analysis was prespecified (Supplementary Data S1), although some variables were deemed unsuitable for use due to high levels of missingness (social class, marital status, education level). All analyses were performed using Stata/IC 15.1.

Descriptive statistics were calculated to compare patients with DAS28<3.2 against those with DAS28≥3.2 values at five and 10 years. For continuous variables, medians and interquartile ranges (IQR) were used and for binary/ordinal, frequencies and proportions. Odds ratios were obtained using binary logistic regression models, separately for five and 10 years, to compare the groups for each covariate under consideration. Five-year outcomes were not used in 10-year models, since only baseline factors were of interest. Variables with a likelihood ratio test p-value <0.2 were tested in multivariable analysis.

Models for five-year and 10-year DAS28 were built using forward selection based on clinical relevance and statistical robustness to obtain odds ratios (ORs) and 95% confidence intervals (CIs). Age at onset, gender and RDCI were included a priori. Variables were added in order of the smallest likelihood ratio test p-values (up to p<0.05) and retained in subsequent steps even if their p-values were increased.

Two interaction terms were deemed necessary for exploration a priori; smoking status with education level and smoking with presence of erosions. After obtaining final models, inclusion of interaction terms was tested in exploratory analyses. All possible interactions were considered, using each covariate in turn. This was done by using the selected covariate in interaction with every other covariate already in the model. Those terms with Wald test values of p<0.15 were added to the original model and assessed for inclusion with likelihood ratio tests. Stratified models were fitted to the data if appropriate, using complete cases.

Once model building was complete, multicollinearity between covariates was assessed, excluding variables with centred variance inflation factors >10. After models were determined using complete cases, multiple imputation by chained equations, performed with 20 imputations, was used. Any missing baseline variables selected in the complete case model, as well as missing DAS28 outcomes, were imputed. Baseline and five-year values of clinical covariates (measured annually) were included in the imputation model for the five-year DAS28 model, including RDCI, similarly baseline and 10-year values in the 10-year DAS28 model, if available. Patients who had died before year five/10
respectively were not included in imputation models (so results should be interpreted as conditional on remaining alive). Continuous variables were imputed using linear regression, binary with logistic regression and ordinal using predictive mean matching.

Sensitivity analyses were performed to determine if any differences were seen depending on the method used to calculate DAS28 scores, described in more detail in Supplementary Data S3.1. A further sensitivity analysis involved fitting a model using all candidate predictors with imputed data is described in Supplementary Data S3.2; this was to mitigate against potential downward bias resulting from the forward selection process. Additionally, any component part of baseline RDCI found to be associated with DAS28 at either timepoint in univariable analysis (p<0.2) was used in place of baseline RDCI in final models to ascertain any relationship with the outcome. Details of this process are described in Supplementary Data S3.3.

**Results**

**Cohort demographics**
There were 2,701 patients recruited, mean age 56.1 years, 66.9% female, with median time of 6 months from disease onset to first outpatient appointment. By year five, 1,718 (63.4%) patients remained enrolled, 185 (6.8%) were known to have died and a further 798 (29.5%) no longer in follow-up. At year 10, 820 (30.4%) were enrolled, 454 (16.8%) had died and 1,427 (52.8%) were not in follow-up. Completeness and availability of patient data are illustrated in the flowchart at Supplementary Data S4. Comparisons between those still enrolled and those no longer followed up are shown in Supplementary Tables S1 and S2. Those still enrolled at five years were younger at disease onset than those no longer followed up (median 55.5 years, IQR 45 to 65 compared to 62, IQR 50 to 70).

**Outcome: Five-Year DAS28**
DAS28 scores were available for 1,354 patients at five years; 808 (59.7%) of patients had values ≥3.2. As shown in Table 1, a higher proportion of patients with DAS28≥3.2 were female (71.9% compared to 60.1%) and slightly fewer of white ethnicity (96.3% vs 98.7%). More patients with DAS28≥3.2 were in the worst deprivation quintile at baseline (14.7% vs 8.4%) and fewer in the least deprived quintile (25.4% vs 30.3%). Of those with DAS28≥3.2, 73.3% had no baseline comorbidities compared to 75.6% with DAS28<3.2. Baseline HAQ-DI was higher amongst the DAS28≥3.2 group (median 1.13, IQR 0.63 to 1.75 compared to median 0.88, IQR 0.38 to 1.38).
Univariable analyses are shown in Supplementary Table S3. No association was seen between RDCI, or any components, and five-year DAS28. Age at disease onset, being female or of minority ethnicity were associated with DAS28≥3.2, each with p<0.2. The odds of DAS28≥3.2 increased in the more deprived compared to the least deprived group. Elevated baseline HAQ-DI, DAS28, BMI and ESR, presence of erosions, being seropositive and low haemoglobin, were each associated with DAS28≥3.2 at five years, p<0.2.

Multivariable Analyses

Results of multivariable analyses are shown in Table 2. Baseline RDCI was not found to be associated with DAS28 at five years (OR 1.05, 95% CI 0.91 to 1.22). Being female increased odds of DAS28≥3.2 by nearly 50% compared to male (OR 1.47, 95% CI 1.17 to 1.85). Baseline HAQ-DI and being seropositive were each associated with increased odds of DAS28≥3.2 (OR 1.50, 95% CI 1.28 to 1.76, and OR 1.33, 95% CI 1.05 to 1.69 respectively). Each increase in IMD quintile, relating to less deprivation at baseline, was associated with reduced odds of DAS28≥3.2 (OR 0.88, 95% CI 0.81 to 0.95). Although ESR featured in the model, its effect on DAS28≥3.2 odds was minimal (OR 1.01, 95% CI 1.00 to 1.01). No evidence was found to include interaction terms in the model; further details in Supplementary Data S2).

Outcome: 10-Year DAS28

At 10 years, DAS28 values were available for 651 patients. A higher proportion had DAS28≥3.2 values at 10 years than at five (65.6% compared to 59.7%). More females had 10-year DAS28≥3.2 (74.5% vs 58.5%) and more patients with DAS28≥3.2 had positive serostatus (77.0% vs 69.5%). Differences in DAS28 according to baseline deprivation remained, with DAS28≥3.2 consisting of a higher proportion of the most deprived quintile than DAS<3.2 (17.1% vs 10.0%), and conversely fewer in the least deprived group (29.3% vs 34.1%).

Univariable analyses (Supplementary Table S3) found no association between RDCI and 10-year DAS28 outcome. A statistically significant difference was seen for depression, ulcer/stomach disease and diabetes groups, but with numbers too small to enable further analysis. As at five years, being female, of minority ethnicity and from a deprived area were associated with increased odds of DAS28≥3.2 (p<0.2). Baseline HAQ-DI, DAS28, BMI, presence of erosions, haemoglobin, being seropositive and ESR were also associated with DAS28≥3.2 (p<0.2).

Multivariable Analyses
In multivariable analyses (table 2), no association was seen between baseline RDCI and 10-year DAS28≥3.2 (OR 0.99, 95% CI 0.75 to 1.31). Baseline HAQ-DI had a smaller effect than seen at five years (OR 1.34, 95% CI 1.06 to 1.69) but remained associated. Conversely, being female saw a larger effect on DAS28≥3.2 odds at 10 years (OR 1.81, 95% CI 1.22 to 2.68). Again, baseline IMD was associated, with each increase in quintile (less deprived) associated with lower odds of DAS28≥3.2 (OR 0.82, 95% CI 0.72 to 0.93). Although seropositive status was associated in complete case analysis, the association was reduced in the imputed model (OR 1.49, 95% CI 0.96 to 2.32). Presence of erosions at baseline was associated with 10-year DAS28≥3.2, in contrast to five years (OR 1.82, 95% CI 1.16 to 2.85). Again, there was no evidence to include interaction terms in the model (Supplementary Data S2).

**Sensitivity Analyses**

There was no evidence to support the exclusion of patients whose DAS28 scores were calculated using CRP instead of ESR (Supplementary Data S3.1 and Supplementary Table S4). Modelling with all candidate predictors had minimal impact on odds ratios for five- or 10-year DAS28 (Supplementary Data S3.2 and Supplementary Table S5). No association was found between any RDCI component and DAS28 at five years, with too few cases to consider separate conditions at 10 years (Supplementary Data S3.3 and Supplementary Table S6). Fitting models with baseline DAS28 in place of ESR resulted in issues with multicollinearity and had little impact on the estimates for other covariates (Supplementary Table S7).

Models fitted separately for each cohort had minimal effect on odds ratios; although the effect of baseline RDCI on five-year DAS28 was slightly greater within the ERAN cohort than ERAS, this remained statistically insignificant (data not shown). As DAS28 data were only available for 1,354 participants at five years and 651 at 10 years, outcome values were imputed for 1,434 (57.0%) and 1,596 (71.0%) of the combined cohorts respectively. Fitting models using only those observed had minimal effect on estimated odds ratios, although presence of erosions at baseline was not found to be significant amongst those with observed 10-year DAS28 scores in contrast to the imputed values (data not shown).

**Discussion**

This study found that baseline comorbidity burden, measured with RDCI, was not associated with DAS28 at either five or 10 years. Age at disease onset had no effect on disease activity, but other sociodemographic factors, namely female gender and worse baseline deprivation were each associated with increased odds of DAS28≥3.2 at both five and 10 years, in line with findings from
other studies(27). Although worse deprivation is known to be associated with poorer HAQ-DI scores, little evidence exists of a relationship with disease activity to date(28). Amongst clinical factors tested, being seropositive was associated with five-year DAS28 and having erosions with 10-year DAS28, as also supported by other literature(29,30). Poorer baseline HAQ-DI was found to associate with DAS28 at both timepoints, again in line with existing evidence(31).

Although it is known that a relationship exists between health in early stages of RA and later disease activity(32), this study has attempted to establish a link between comorbidities present at baseline and worse disease outcomes. RA patients are known to be at increased risk of cardiovascular disease(33), malignancies, infections, gastrointestinal diseases, osteoporosis and depression(34), but with little research undertaken to assess the relationship between comorbidities at onset of RA and subsequent disease activity. Those living in areas of the UK with more deprivation are known to have poorer health outcomes(35) and this study has confirmed that this is also true in the case of people living with RA. This highlights the value of considering the whole picture when seeing a patient in clinic for the first time, including sociodemographic as well as clinical factors.

Previous research from our group showed that baseline comorbidities were not associated with later functional outcomes(36), but that sociodemographic factors (increased age at disease onset, female gender and minority ethnicity) were associated with increased odds of HAQ-DI≥1.5 at both five and 10 years. Additionally, worse deprivation at baseline was associated with 10-year HAQ-DI≥1.5. Insights from these works will be used to inform risk-stratification of patients and specific pathways to enhance the care of patients with RA. Sociodemographic factors at baseline are informative and bring another important dimension that deserves attention, since they are unlikely to vary in subsequent years, at least not as much as clinical measures. The role in risk stratification early on, is therefore crucial.

This study is unique in several ways. Firstly, in the use of two large early RA cohorts, long in follow up, rich in clinical, comorbidity, sociodemographic data and designed such that they could be analysed collectively. Secondly, in studying both biological and non-biological factors, recognising the impact that comorbidities have on T2T goals. Finally, in investigating sociodemographic factors, particularly deprivation, adding a further dimension and allowing the study of associations with other clinical parameters and the influence on patient outcomes. However, there are also some limitations, including the age of these cohorts, during which time treatment may not have reflected modern approaches, limiting the generalisability of these findings. DAS28 values were available for
just 1,354 participants at five years and 651 at 10 years, with multiple imputation applied to address this. It was assumed that patients without follow up data were missing at random, but it is possible that they could have had higher disease activity than those remaining in the study. Although similar results were seen amongst those with measured outcome values, consideration should be given of the high proportion of imputed values when interpreting these results. Few patients who identified as belonging to a minority ethnicity group participated, such that responses could not be meaningfully analysed to represent different communities. Links between ethnicity and deprivation within the UK are well-documented(37) but could not be further explored in this dataset. A further limitation is the number of hypotheses tested and lack of multiple testing correction. It is possible that some significant results are false positives, and results should be interpreted as exploratory rather than confirmatory. Overall, despite the use of historical RA inception cohorts, insights from this study are relevant even in modern rheumatology practice particularly with the added dimension of the potential role of social factors in risk-stratification and in informing a more holistic management of patients with RA, at first presentation to rheumatology.

In conclusion, this study found no association between baseline comorbidities and DAS28 outcome at five or 10 years in RA patients. However, some sociodemographic and clinical factors recorded at baseline play a role in disease activity in later years, in models adjusted for comorbidity burden. These findings build upon those relating to functional outcome and will accelerate the identification at initial rheumatology appointment of RA patients at increased risk of worse outcomes. Our findings support that tailored intervention can be implemented in a timely manner, in line with national and international recommendations, potentially limiting permanent joint damage and disability and improving quality of life for individuals.

**Funding**
This work was supported by the British Medical Association (BMA) as part of the Doris Hillier award received by EN.

**Conflicts of Interest**
The authors declare no conflicts of interest related to this work.

**Ethics Approval**
Obtained from East Hertfordshire local research ethics committee (ERAS) and the Trent research ethics committee (ERAN).
Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.
References


4. European Alliance of Associations for Rheumatology (EULAR) [Internet]. Available from: https://www.eular.org/eular_about.cfm


26. NRAS. The relevance and importance of DAS28 in your treatment pathway [Internet]. Available from: https://nras.org.uk/resource/the-das28-score/


