Article:
Dagliati, A, Plant, D, Nair, N et al. (10 more authors) (2020) Latent class trajectory modelling of 2-components-DAS28 identifies multiple rheumatoid arthritis phenotypes of response to biologic disease modifying anti-rheumatic drugs. Arthritis & Rheumatology, 72 (10). pp. 1632-1642. ISSN 2326-5191

https://doi.org/10.1002/art.41379

DR. ARIANNA DAGLIATI (Orcid ID : 0000-0002-5041-0409)
DR. MEGHNA JANI (Orcid ID : 0000-0002-1487-277X)
PROF. ANTHONY GERARD WILSON (Orcid ID : 0000-0003-4855-3926)

Article type : Full Length

Latent class trajectory modelling of 2-components-DAS28 identifies multiple rheumatoid arthritis phenotypes of response to biologic disease modifying anti-rheumatic drugs.

Arianna Dagliati, PhD[1,2‡], Darren Plant, PhD[3,4‡], Nisha Nair, PhD[3], Meghna Jani, PhD[6], Beatrice Amico, MS[9], Niels Peek, PhD[1,4], Ann W Morgan, PhD[7], John Isaacs, PhD[5], Anthony G Wilson, PhD[8], Kimme L Hyrich, PhD[4,6], BRAGGSS[10], Nophar Geifman, PhD[1,2□] and Anne Barton, PhD[3,4,□§]

1. Centre for Health Informatics, The University of Manchester, Manchester, UK.
2. The Manchester Molecular Pathology Innovation Centre, The University of Manchester, UK.
3. Versus Arthritis Centre for Genetics and Genomics, The University of Manchester, Manchester, UK.
4. NIHR Manchester BRC, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK.
5. Translational and Clinical Research Institute, Newcastle University and Musculoskeletal Directorate, Newcastle upon Tyne Hospitals NHS Foundation Trust
6. Versus Arthritis Centre for Epidemiology, The University of Manchester, Manchester, UK
7. School of Medicine, University of Leeds and NIHR Leeds BRC, Leeds Teaching Hospitals NHS Trust
8. Conway Institute, Centre for Arthritis Research, University College Dublin, Dublin, Ireland
9. Department of Compute Science, University of Verona, Italy
10. BRAGGSS, See Supplementary

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ART.41379

‡ Equal first author, AD and DP shared first authorship, both authors contributed equally

□ Equal last author, NG and AB shared last authorship, both authors contributed equally

§ Corresponding author:

Anne Barton, Arthritis Research UK Centre for Genetics and Genomics, Centre for Musculoskeletal Research,
Manchester Academic Health Science Centre,
The University of Manchester, Manchester M13 9PT, UK.
E-mail: anne.barton@manchester.ac.uk

Competing interests

The authors have no competing interests to report

Funding information

MATURA was jointly funded by MRC and Versus Arthritis (grant ref MR/K015346/1), NIHR Manchester BRC, MRC Molecular Pathology Node MMPathic (grant ref MR/N00583X/1), NIHR Newcastle BRC and Versus Arthritis (grant refs 21754 and 22072)

Key Words

Rheumatoid Arthritis, Disease Activity, DMARDs (biologic), Treatment, Longitudinal Data.

Abstract

Objectives To determine whether using a re-weighted disease activity score that better reflects joint synovitis, the 2-component DAS28 (2C-DAS28), based on 28-swollen joint count and C-reactive protein, produces more clinically relevant treatment outcome trajectories compared with the standard 4 component DAS28 (4C-DAS28).

Methods Latent class mixed modelling (LCMM) of response to biologic treatment was applied to 2,991 patients with RA about to commence bDMARD treatment from the Biologics in Rheumatoid Arthritis Genetics and Genomics Study Syndicate (BRAGGSS) cohort using both 4C-DAS28 and 2C-DAS28 as outcomes. Identified patient groups with similar trajectories were compared in terms of pre-treatment baseline characteristics – including disability, comorbidities - and follow-up characteristics – including anti-drug antibody (ADAb) events, adherence to treatments and blood drug levels, We compared the trajectories obtained using the 4C and 2C scores to determine which characteristics were better captured by the 2CDAS28 compared with the 4C-DAS28 trajectories.

Results. Using the 4C-DAS28 we identified 3 trajectory groups, which is consistent with previous reports. We show that the 4C-DAS28 captures information relating to depression. Using the 2C-DAS28, 7 trajectory groups were identified; among them, distinct groups of non-responders had a higher incidence of respiratory comorbidities and a higher proportion of ADAb events. We also identified a group of participants for which the 2C-DAS28 disease activity score remained relatively low and which was enriched for patients who were non-adherent to treatment. This highlights the utility of both the 4C and 2C-DAS28 for monitoring different components of disease activity.

Conclusions Here we show that the 2C-DAS28 modified disease activity score defines important biological and clinical phenotypes associated with treatment outcome in RA and characterizes important underlying response mechanisms to biologic drugs.

Introduction

Complexity and variability of patient trajectories in response to treatment poses significant challenges in many medical fields[1–4]. A current focus of medical research is facilitating a shift away from a 'one-size-fit-all' to precision medicine approaches, aimed at identifying narrower patient subgroups who would benefit from tailored interventions. However, the high variability in response to therapeutic agents over time[5–7] makes prediction of response challenging.

Among subgroup discovery approaches from longitudinal data, latent class mixed modelling (LCMM) is gaining increasing popularity for identifying homogeneous subgroups[8–12]. LCMM has been applied to study disease progression and response to treatments in psychiatry, nephropathy, neurology[13–15] and early arthritis[16]. In rheumatoid arthritis (RA), LCMM has been used to identify trajectories of disability progression following diagnosis of RA[17]; disease activity using the disease activity score (DAS28)[18–20] for Treat-to-Target strategies[18], response to biologic agents[19], and prediction of remission in people treated with disease modifying anti-rheumatic drugs (DMARDs)[20]. All of these studies used the 4-componenet (4C) disease activity score measured in 28 joints (4C-DAS28) and identified three broad trajectories.

It has previously been shown that the 4C-DAS28 correlates poorly with synovitis[21]. Given that the drugs used to treat RA aim to reduce synovial inflammation and prevent joint damage, it is unclear how useful the 4C-DAS28 trajectories will be as outcomes for precision medicine studies, which aim to identify biomarkers that are predictive of therapeutic response. To address this, a revised composite disease activity score, including only the 28-swollen joint count and C-reactive protein[21] (2C-DAS28), has been developed and shown to correlate better with ultrasound-detected synovial inflammation compared to the original 4C-DAS28[21].

This study aims to apply the LCMM approach in a prospective, longitudinal cohort of patients starting biologic (b)DMARDs using the original 4C-DAS28 and the 2C-DAS28 and compare whether the identified trajectories have clinical relevance by exploring their association with comorbidities, treatment adherence and drug immunogenicity, all of which have previously been reported to associate with bDMARD treatment response[22,23].

Materials and Methods

Study Participant and Data

We utilized longitudinal data (November 2008/January 2018) from patients recruited into Biologics in Rheumatoid Arthritis Genetics and Genomics Study Syndicate (BRAGGSS), a UK observational multi-centre study established to study predictors of treatment-response to bDMARDs[24]. Consenting participants were eligible if they had a consultant diagnosis of RA, were over 18 years of age and about to commence bDMARD treatments, namely TNF, B-cell-Inhibitor (anti-CD20), IL6R or T-cell-Inhibitor (CTLA4). Information regarding the combination with conventional DMARD (cDMARD) e.g. methotrexate (MTX), were recorded. During the observational study a minority of participants suspended, changed, or re-started bDMARD therapy. These subjects were included in the analyses and information about previous treatments with bDMARD was recorded. Following the pre-treatment baseline sample and data collection, participants were observed for 12 months with follow-ups at 3, 6 and 12 months.

Contributing patients provided written informed consent, and the BRAGGSS study is ethically approved (COREC 04/Q1403/37).

The primary outcome measure to model response-trajectories was the disease activity score based on 2-Components-DAS28 (2C-DAS28)[21], computed as:

$$\text{2C-DAS28} = \text{sqrt}(\text{Swollen28}) + (0.6 * \ln(\text{CRP}+1)) \quad \text{(eq.1)}$$

4-Component-DAS28 (4C-DAS28) and response criteria were calculated according to standard methodology[25].

Variables recorded at baseline included: Age, sex, time since diagnosis, body mass index (BMI), smoking status (if the subject had ever smoked), functional severity measured by the Health Assessment Questionnaire[26] (HAQ), and consultant-reported comorbidities.

DAS28 components (Tender28, Swollen28, Patient global assessment of wellbeing measured on a 100 mm Visual Analogue Scale (VAS), serum CRP level) were measured at baseline (i.e. before commencing any bDMARD treatment) and at each follow-up.

For participants treated with TNF-inhibitors, anti-drug antibodies (ADAb) and serum drug levels were measured in serum samples collected at follow-up visits.

We included participants with at least one 2C-DAS28 measurement after baseline, including all measurements until the last registered follow-up. Unlike previous studies [27,28], we did not exclude subjects on the basis of the number of measures to avoid possible selection biases, and we also included subjects who had received biological treatment in the past. Participants who switched bDMARD were right censored when the change occurred.

Missing data arose from non-completion of demographic characteristics and missing HAQ scores. Due to the nature of the variables, we assumed they were missing at random. As in [29,30], we compared different imputations strategies (see Appendix paragraph "Imputation Strategy") and implemented the method with the best performance. Data were imputed with a Random Forest approach with 100 trees and a maximum of 100 iterations.

Latent Class Analysis

We applied LCMM to identify patient subgroups with distinct responses to bDMARDs over time. Previous stratification studies in RA using LCMM focused on 4C-DAS28 as the measure of disease activity and response to biologic treatment[18–20]; those analyses identified three groups of responders: Rapid, Gradual and Inadequate. Here we specified two linear mixed-effects models: (i) using 4C-DAS28 as in previous studies and (ii) using 2C-DAS28 as the dependent variable.

Mixed effects were used to account for the likely correlation of repeated measurements and included a random intercept for each individual. We fitted the model through the 'lcmm' function of the R package lcmm[31]. We followed the framework from Lennon et al. [32] that includes scoping model definition, refinement of the number of classes and model structure, model adequacy assessment and clinical plausibility, graphical presentations and sensitivity analysis (detailed in the Appendix).

Models were adjusted for age, time since diagnosis, gender, BMI, HAQ, bDMARD therapy, concomitant treatment with MTX, and previous treatment with biologic drugs. We developed scoping models for 1 to 10 classes to determine the optimal number of classes based on the Bayesian Information Criteria (BIC) and further refined the models comparing linear quadratic and cubic specification of time (as reported in the Appendix, steps 3 and 4 of the LCMM Framework).

To evaluate whether results were driven by differential relative effectiveness of therapies, in sensitivity analyses we repeated the analysis in participants treated with TNF-inhibitors only and in participants treated with any bDMARD in combination with MTX and compared the discovered latent classes in terms of baseline characteristics and response variables.

Statistical description and comparison of latent classes

Discovered latent classes were described and compared in terms of clinical characteristics, comorbidities, adherence to treatments, ADAb positivity and non-trough blood drug levels.

Intra-individual variation in treatment response

To describe intra-individual variation in treatment-response over time, we computed slopes as the difference between consecutive measurements divided by the time between the measurements (Supplementary Eq.1). Higher slope values indicate deterioration of the patient's condition, whilst negative values indicate improvements. Time to non-response was defined as a switch from a negative to a positive value of the slope (i.e. when the slopes start to increase). Intercept values and slopes were compared between classes using ANOVA tests. Time to non-response was compared in non-responders using Kaplan Meier visualization.

Demographic and clinical characteristics

Clinical characteristics were compared among latent classes. Categorical characteristics (gender, pre-treatment smoking habits, MTX co-therapy and change of bDMARDs) were compared using chi-square tests with Bonferroni-Holm corrections, whilst continuous variables (age, disease duration, HAQ, BMI, DAS28 components) were compared using ANOVA with post-hoc Tukey tests.

Association of latent classes with comorbidities and therapy (TNF-inhibitors, B-cell-inhibitor, IL6R-inhibitor and T-cell-inhibitor) was explored using logistic regression, where the latent class assignment was the regression predictor, and least square means for multiple comparisons with Tukey's adjustment (from the R lsmeans[33] package). The effect of respiratory comorbidities on latent classes was assessed with stratified analyses in smokers (current or past) and non-smokers (never).

Adherence, drugs levels and ADAb positivity

We studied the associations of adherence, drugs levels, and ADAb positivity with the identified latent classes. Drug-level and antidrug antibodies values have been previously shown [23] to be associted with response of bDMARD. Self-reported adherence, was recorded at 3 and 6 months following the start of therapy[34]; adherent patients were defined as those who took the dose on the day agreed with their healthcare team or no more than one day before or after whilst non-adherent patients were defined as those who self-reported either missing the dose completely or taking it more than one day before or after the agreed day.

The correlation between drug levels and latent classes was investigated in terms of subject level relative change at the next follow-up. We created matched samples with propensity scoring, including BMI (as in[35]), age, sex as potential confounders without interactions and drug levels changes (increasing or decreasing levels) as a binary outcome. Chi-square tests were then applied to study differences in drug level changes among latent classes.

To study ADAb development, we used a threshold of 12 AU/ml as a positive event [23]. Heterogeneities of latent classes across the total cohort and the subset with ADAb information were determined by Cochrane's Q-statistic. ADAb development event was compared for latent classes using Kaplan Meier visualization and Cox proportional hazards regression model adjusting for age, sex, BMI and concomitant treatment with MTX.

Analyses were computed using R version 3.2.3. Results are presented as the main effect with 95% confidence intervals and 5% significance level for main inferences.

Results

Our dataset included 2,991 participants with 7,567 DAS28 measurements after baseline, excluding the baseline measure. The number of follow-ups per participant ranged from 1 to 3 with the following proportions: 1647 (55.06%) subjects had 1 follow-up, 661 (22.09%) had 2 follow-ups measures and 503 (16.82%) had 3 follow-ups. The first follow-up measurement took place on average 124 (SD 51) days after baseline, the second 217 (SD 51) days, and third 365 (SD 50) days. Table 1 reports baseline characteristics, including percentages of missing observations.

4C-DAS28 trajectories and comparison with previous studies

We reproduced a 3-trajectory model using 4C-DAS28, with three classes producing the lowest BIC and highest assignment posterior probability compared to models ranging from 1 to 10 classes (Supplementary reports BIC for each model in Fig.18 and misclassification rates in Table13).

Although inclusion criteria for the studies were different[18,20], and mean 4C-DAS28 at baseline higher[19], response patterns were qualitatively comparable to previous studies. We identified: a Rapid (n=2,003, 66.98%) group, which improved quickly in the first observation period then stabilized or showed slight increases in disease activity, a Gradual (n=919, 30.73%) group, with slower but consistent decrease in disease activity and Poor (n=68, 2.27%) responders (Supplementary reports the discovered trajectories in Fig.19 and their comparison with previous studies in Table14).

4C-DAS28 classes were compared with response patterns calculated with EULAR criteria at different time points. As expected, grouping by LCMM and EULAR revealed proportional differences with the majority of patients in the three LCMM classes grouped as either non- or intermediate-responders using the EULAR criteria. The comparison of EULAR responses between the latent classes confirmed the findings: EULAR Non-response was reached by the majority of Poor responders at the third follow-up, Gradual responders were enriched for EULAR Moderate Responders, with more than half of the population being classified as moderate responders by the third follow-up, whilst the majority of Rapid responders were classified as EULAR Good Responders by the third follow-up. (Supplementary reports the number and proportion of subjects classified via EULAR criteria in Tab.15 and Fig.20).

Comparison of the 4C-DAS28 latent trajectories identified differences related to depression, extracted from reported comorbidities, with borderline statistical evidence for higher rates in Gradual Responders compared to Rapid responders (Rapid, n=398, 20.4%; Gradual, n=214, 24.2%; OR=1.24, RR=1.05, p=0.08), as reported in the supplementary Fig.21, Tab.14.

We found no significant correlation between drug level variation and response-trajectories (p=0.40) using the 4C-DAS28 to define the trajectories. ADAb were available for 475 patients receiving Adalimumab or Certolizumab, whose disease activity classes were similarly distributed as for the whole cohort (Supplementary Tab.12). ADAb positivity was higher in Gradual

Responders compared to Rapid responders (beta=1.79, 95% CI:1.30-2.46, p-value=0.001) (Supplementary Fig.23-24).

2C-DAS28 latent class analysis identifies 7 trajectories of response

2C-DAS28 based LCMM analysis (complete results are reported in the Appendix paragraph "Latent Class Analysis") identified 7 trajectories, which were arbitrarily labelled according to the score at pre-treatment and change in score over time as: 2 groups of Good Responders, 2 Gradual Responders, 2 Secondary Non-Responders and 1 Low Disease Activity (Figure 1). Individual DAS28 component trajectories in the discovered latent classes are shown in Supplementary Fig.12.

- Good responders (Figure 1, green trajectories) demonstrated fast improvement in disease activity in the first 3 months, followed by maintenance of good response. The first group of good responders (Good 1; n=395 subjects, 13.2%) and the second (Good 2; n=1840, 61.6%) differed significantly (p-value<2.2e-16) in baseline 2C-DAS; with the first showing higher values (mean=5.85, SD=0.57) compared to the second group (mean=3.94, SD=0.79).
- Gradual Responders (Figure 1, blue trajectories) demonstrated a continuous but gradual reduction in disease activity. The two classes had significantly (p-value<2.2e-16) different 2C-DAS28 baseline values, with the first group (Gradual 1; n=65, 2.2%) having higher values (mean=6.33, SD=0.79) than the second group (Gradual 2; n=506, 16.9%, mean=4.24, SD=0.85).
- Secondary Non-Responders (Figure 1, red and orange trajectories) showed a fast improvement in disease activity scores followed by rapid deterioration. This class encompasses two subgroups: Early Secondary Non-Responders (n=33, 1.1%) and Late Secondary Non-Responders (n=64, 2.1%). Early Non-Responders had significantly faster slopes of deterioration (p-value = 0.044, diff=0.072, lower= 0.001, upper=0.144) compared to Late Non-Responders.
- Low disease activity patients (Figure 1, yellow trajectories) (n=88, 2.9%) demonstrated the lowest disease activity at baseline (mean 2C-DAS28=2.24, SD=1.09), which increased modestly during follow-up. Initial levels of disease activity measured using 4C-DAS28

were driven by Tender28 and VAS components, while Swollen28 and CRP showed a consistent increase in the follow-up period (Figure 2).

Latent classes were described in term of slopes (Supplementary Fig.10 and Table 6 reports the comparison of slopes at different follow-ups). When we compared Secondary Non-Responders in terms of 2C-DAS28 slopes and time to non-response, we found that Early Secondary Non-Responders lost response significantly (p-value=0.005) earlier (on average 212 days after the baseline visit, SD=28 days) than Late Secondary Non-Responders (265 days, SD=73 days), Supplementary Fig.11 shows the Kaplan Meier curves of Secondary Non-Responders loss of responses.

Individuals in the low disease activity group showed modest improvements in Tender28 and VAS counts over the observation period but deterioration in the Swollen28 score and CRP (Figure 2, panel B). The majority of patients in the low disease activity group were classed as non-responders (79% by 6-months) using 4C-DAS28 derived EULAR response criteria (Supplementary Tab.8-9 reports the response of Low disease activity participants in terms of EULAR responses and individual DAS28 subcomponents). Importantly, the low disease activity is the group that includes the majority of subjects who had received other bDMARD treatments in the past (36.36%).

Sensitivity analyses showed that in TNF-inhibitor treated subjects (2,151 participants), the best model identified 6 classes (Supplementary Fig.7 reports the discovered 2C-DAS28 trajectories of TNF-inhibitor treated subjects) with absence of the gradual responder class with low initial 2C-DAS28 value. In patients treated concomitantly with TNF-inhibitors and MTX (1,878 participants), the best fitted model included 7 classes (Supplementary Fig.4 reports the discovered 2C-DAS28 trajectories of MTX treated subjects), with trajectories comparable to the whole cohort model.

2C-DAS28 response trajectories are associated with different baseline characteristics

LCMM analysis revealed significant differences among classes in: time from diagnosis, HAQ scores and number of comorbidities. Basic statistics and p-values are reported in Table 2. Further details are in the Supplementary section Baseline characteristics.

Figure 3 illustrates comorbidities prevalence in LCMM classes. Early secondary Non-Responders had a higher prevalence of chronic bronchitis and emphysema compared with Good (Good 1: p-value=0.005, OR=5.38 and, Good 2: p-value=0.0005, OR=5.85) and Gradual Responders (Gradual 2; p-value=0.003, OR=5.54). The stratified analysis of correlation between being an Early secondary Non-Responder and having chronic bronchitis/emphysema found significant effects in ever-smokers (OR=2.49 (1.14, 5.42)) and never-smokers (OR=4.87 (1.06, 22.32)). Low Disease Activity patients had a higher prevalence of asthma compared with Good Responders (Good 1, p-value=0.029, OR=2.53).

2C-DAS28 response trajectories are associated with Adherence, Drugs levels and ADAb.

Self-reported adherence measures, collected at follow-up visits, were available for 1,528 subjects, 51% of the entire cohort. 2C-DAS28 slopes differed (p-value=0.002) and decreased faster in adherent (mean=-0.00128, SD=0.0581) than in non-adherent (mean=-0.0120, SD=0.103) participants. Adherence patterns (Supplementary Fig.16 illustrates the distribution of adherence in latent classes) indicate that the Low disease activity group, which are only identified using 2C-DAS28, were the least adherent group.

Differences in adherence were confirmed by blood drug levels changes (Supplementary Fig.17 and Tab.11 reports the drugs level at different follow-ups stratified by type of bDMARD). Significant correlation between drug level changes at the second follow-up and 2C-DAS28 response-trajectories (p-value = 0.03) were observed in the matched data set (n=524 subjects).

Regarding ADAb positivity, the highest proportion of ADAb positive patients were observed in the Early Secondary Non-Responders, but high ADAb titres were also observed in Gradual Responders (Figure 4 A). Cox regression models (Figure 4 B), including the effect of MTX co-treatment, revealed different risks for ADAb (Beta=6.06, 95% CI:2.57-14.27, p-value=3.71e-05 for Late Secondary non-responders compared to Good 1; Beta=1.79, 95% CI:1.00-3.22, p-value=6.06, 95% CI:2.57-14.27 for Gradual 2 compared to Good 1).

Discussion

We have found that using the 2C-DAS28 score as an outcome measure reveals more subgroups of patients compared to the 4C-DAS28 and the trajectories identified are clinically meaningful.

Whilst the 4C-DAS28 remain an essential tool to assess RA progression in clinical practice, the three trajectories identified using the 4C-DAS28, which can be also interpreted in terms of EULAR responses, may not provide sufficient granularity to help unpick important biological mechanisms underpinning inflammation.

Strengths of the study are that it is the first and largest study, based on real-world patients, to explore patient-centred factors to trajectories of drug response; the response-trajectories based on 2C-DAS28 were identified using a structured methodological framework[32]; all classes of licenced bDMARD treatment regimens were considered and included in longitudinal data models; the clinical relevance of the identified subgroups were compared using outcomes relevant to patients.

We found that the 7 subgroups demonstrated different degrees of treatment response, mean patterns of change in disease activity, and different clinical characteristics. For example, drug levels are better tracked through the 2C-DAS28 seven trajectories than the three 4C-DAS28 trajectories. Furthermore, ADAb positivity was enriched in two of the 2C-DAS28 Non-Responder groups, indicating these trajectories are biologically relevant and more informative than those identified by 4C-DAS28.

Another potentially clinically useful subgroup identified using the 2C-DAS28 was the low disease activity group, where conventional DAS28 scores were driven by the tender joint count and VAS. This group were less likely to respond to treatment (79% classed as EULAR non-responders by 6 months) and were less likely to adhere to treatment. It is also interesting to note that previous non-adherence, together with an increased resistance to therapy, might be related to a higher proportion of subjects receiving previous bDMARDs treatment (nearly the 37% within the group). Using biologic drugs in this subgroup is unlikely to be clinically or cost-effective and biomarkers to identify such patients would be useful.

While 4C-DAS28 is an important holistic assessment, which better captures depression, 2C-DAS28 appears better able to capture biological comorbidities, especially respiratory phenotypes. As previously demonstrated, ADAb concentrations and treatment adherence remained the most important predictors of drug levels over time[22]. The 2C-DAS28 trajectories capture these

characteristics and confirm previous findings that drug levels, adherence and ADAb positivity correlate with speed and sustainability of response.

A number of limitations should be considered. First, while the 2C-DAS28 score was developed to correlate better with synovitis and validated by demonstrating increased association with radiographic progression in early RA, its weighting remains to be optimized in established RA and RA patients receiving bDMARDs and the implications for treatment decisions remain to be established. Second, drugs levels and ADAb measures were only available in a subset of participants, even if similarly distributed across the 7 trajectory subgroups. Furthermore, adherence to treatments was self-reported. Third, although we studied baseline comorbidities, adverse events leading to treatment cessation are potential confounders, especially within patients showing high rates of non-adherence but we did not have sufficient information on those events to include in the current analysis. A further limitation relates to the selection of participants in the BRAGGSS study, following UK NICE guideline [36]; thus the presented results are generalizable only to populations whose treatments and clinical characterises are comparable with these guidelines.

A final limitation regards the exploitation of unsupervised approaches and the exploratory hypothesis-generating nature of these approaches. However, we mitigated this effect by using clinicians' experience and the previous literature (i.e. suggesting 3 classes of responders as measured via 4C-DAS28 - which we replicate in this study), following a robust analysis framework (which entailed testing and re-testing of multiple models) and performing sensitivity analyses on two subsets to verify our results.

In research for which the primary interest is characterising improvement in biological changes in synovitis, relying solely on the 4C-DAS28 trajectories may result in misclassification of synovial inflammatory responses. 2C-DAS28 trajectories provide finer granularity and may be a better measure to reveal underlying biology. Whilst there is currently no implemented method to predict class assignment for new patients, research to identify molecular biomarkers that can discriminate between trajectory classes at baseline will enable this. Thus, we suggest the use of 2C-DAS28 score for biomarker discovery studies, which should encompass the understanding of adherence to treatments, immunogenicity to drugs, and general health status, including HAQ scores and comorbidities.

Better defined and physiologically contrasted phenotypes imply better understanding of the underlying responses mechanisms and are the pillar for precision medicine and well-powered biomarker discovery. Previous studies have shown that patients themselves want precision medicine approaches to receive medications that are more likely to work more quickly[37]. Therefore, more robust prediction models for bDMARDs responses should be built on the basis of outcomes reflecting the disease biology, provide evidence of larger variability to biological treatments and allow the identification of essential covariates to build these models. This study shows how 2C-DAS28 trajectories capture these aspects better than traditional outcomes and suggest that a more holistic view of patient care, considering all factors likely to influence response, should be applied to move towards precision medicine approaches.

References

1 Giannoula A, Gutierrez-Sacristán A, Bravo Á, *et al.* Identifying temporal patterns in patient disease trajectories using dynamic time warping: A population-based study. *Sci Rep* 2018;**8**. doi:10.1038/s41598-018-22578-1

2 Jensen AB, Moseley PL, Oprea TI, *et al.* Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nat Commun* Published Online First: 2014. doi:10.1038/ncomms5022

3 Defossez G, Rollet A, Dameron O, *et al.* Temporal representation of care trajectories of cancer patients using data from a regional information system: An application in breast cancer. *BMC Med Inform Decis Mak* Published Online First: 2014. doi:10.1186/1472-6947-14-24

4 Joyce DW, Kehagia AA, Tracy DK, *et al.* Realising stratified psychiatry using multidimensional signatures and trajectories. *J Transl Med* Published Online First: 2017. doi:10.1186/s12967-016-1116-1

5 Edwards RR, Dworkin RH, Turk DC, *et al.* Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain. 2016. doi:10.1097/j.pain.0000000000000602

6 Tenenbaum JD, Avillach P, Benham-Hutchins M, *et al.* An informatics research agenda to support precision medicine: seven key areas. *J Am Med Informatics Assoc* 2016;**1**:ocv213. doi:10.1093/jamia/ocv213

7 Gambhir SS, Ge TJ, Vermesh O, *et al.* Toward achieving precision health. Sci. Transl. Med. 2018. doi:10.1126/scitranslmed.aao3612

8 Geifman N, Lennon H, Peek N. *Patient stratification using longitudinal data - Application of latent class mixed models*. 2018. doi:10.3233/978-1-61499-852-5-176

9 Geifman N, Kennedy RE, Schneider LS, *et al.* Data-driven identification of endophenotypes of Alzheimer's disease progression: Implications for clinical trials and therapeutic interventions. *Alzheimer's Res Ther* 2018;**10**. doi:10.1186/s13195-017-0332-0

10 Vistisen D, Witte DR, Tabák AG, *et al.* Patterns of Obesity Development before the Diagnosis of Type 2 Diabetes: The Whitehall II Cohort Study. *PLoS Med* Published Online

First: 2014. doi:10.1371/journal.pmed.1001602

11 Zheng Y, Song M, Manson JE, *et al.* Group-Based Trajectory of Body Shape from Ages 5 to 55 Years and Cardiometabolic Disease Risk in 2 US Cohorts. *Am J Epidemiol* Published Online First: 2017. doi:10.1093/aje/kwx188

12 Lin H, Turnbull BW, McCulloch CE, *et al.* Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer. *J Am Stat Assoc* Published Online First: 2002. doi:10.1198/016214502753479220

13 Verger P, Mmadi Mrenda B, Cortaredona S, *et al.* Trajectory analysis of anxiolytic dispensing over 10 years among new users aged 50 and older. *Acta Psychiatr Scand* Published Online First: 2018. doi:10.1111/acps.12858

14 Boucquemont J, Loubère L, Metzger M, *et al.* Identifying subgroups of renal function trajectories. *Nephrol Dial Transplant* Published Online First: 2017. doi:10.1093/ndt/gfw380

15 Signori A, Izquierdo G, Lugaresi A, *et al.* Long-term disability trajectories in primary progressive MS patients: A latent class growth analysis. *Mult Scler J* Published Online First: 2018. doi:10.1177/1352458517703800

16 Barnabe C, Sun Y, Boire G, *et al.* Heterogeneous disease trajectories explain variable radiographic, function and quality of life outcomes in the Canadian Early Arthritis Cohort (CATCH). *PLoS One* Published Online First: 2015. doi:10.1371/journal.pone.0135327

17 Norton S, Sacker A, Dixey J, *et al.* Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. *Rheumatology (Oxford)* Published Online First: 2013. doi:10.1093/rheumatology/ket253

18 Siemons L, Ten Klooster PM, Vonkeman HE, *et al.* Distinct trajectories of disease activity over the first year in early rheumatoid arthritis patients following a treat-to-target strategy. *Arthritis Care Res* Published Online First: 2014. doi:10.1002/acr.22175

19 Courvoisier DS, Alpizar-Rodriguez D, Gottenberg JE, *et al.* Rheumatoid Arthritis Patients after Initiation of a New Biologic Agent: Trajectories of Disease Activity in a Large Multinational Cohort Study. *EBioMedicine* Published Online First: 2016. doi:10.1016/j.ebiom.2016.08.024

20 Novel methodology to discern predictors of remission and patterns of disease activity over time using rheumatoid arthritis clinical trials data. *RMD Open* Published Online First: 2018. doi:10.1136/rmdopen-2018-000721

21 Hensor EMA, McKeigue P, Ling SF, *et al.* Validity of a two-component imaging-derived disease activity score for improved assessment of synovitis in early rheumatoid arthritis. *Rheumatology* Published Online First: 2019. doi:10.1093/rheumatology/kez049

22 Jani M, Isaacs JD, Morgan AW, *et al.* High frequency of antidrug antibodies and association of random drug levels with efficacy in certolizumab pegol-treated patients with rheumatoid arthritis: Results from the BRAGGSS cohort. *Ann Rheum Dis* Published Online First: 2017. doi:10.1136/annrheumdis-2015-208849

23 Jani M, Chinoy H, Warren RB, *et al.* Clinical utility of random anti-tumor necrosis factor drug-level testing and measurement of antidrug antibodies on the long-term treatment response in rheumatoid arthritis. *Arthritis Rheumatol* Published Online First: 2015. doi:10.1002/art.39169

24 Cordingley L, Prajapati R, Plant D, *et al.* Impact of psychological factors on subjective disease activity assessments in patients with severe rheumatoid arthritis. *Arthritis Care Res* Published Online First: 2014. doi:10.1002/acr.22249

25 DAS-score. https://www.das-score.nl/das28/en/

26 Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005.

27 Verheij RA, Curcin V, Delaney BC, *et al.* Possible Sources of Bias in Primary Care Electronic Health Record Data Use and Reuse. *J Med Internet Res* Published Online First: 2018. doi:10.2196/jmir.9134

28 Weber GM, Adams WG, Bernstam E V., *et al.* Biases introduced by filtering electronic health records for patients with 'complete data'. *J Am Med Informatics Assoc* Published Online First: 2017. doi:10.1093/jamia/ocx071

29 Dagliati A, Marini S, Sacchi L, *et al.* Machine Learning Methods to Predict Diabetes Complications. *J Diabetes Sci Technol* Published Online First: 2017. doi:10.1177/1932296817706375

30 Celton M, Malpertuy A, Lelandais G, *et al.* Comparative analysis of missing value imputation methods to improve clustering and interpretation of microarray experiments. *BMC Genomics* Published Online First: 2010. doi:10.1186/1471-2164-11-15

31 Proust-Lima C, Philipps V, Liquet B. Estimation of Latent Class Linear Mixed Models : the new package lcmm. *arXiv* 2016;:1–13.

32 Lennon H, Kelly S, Sperrin M, *et al.* Framework to construct and interpret latent class

trajectory modelling. *BMJ Open* 2018;**8**:20683. doi:10.1136/

33 Lenth R. Package ‘ lsmeans ’. *J Stat Softw* Published Online First: 2018. doi:10.1080/00031305.1980.10483031>.NOTE

34 Ling S, Bluett J, Barton A. Prediction of response to methotrexate in rheumatoid arthritis. Expert Rev. Clin. Immunol. 2018. doi:10.1080/1744666X.2018.1465409

35 Ottaviani S, Gardette A, Tubach F, *et al.* Body mass index and response to infliximab in rheumatoid arthritis. *Clin Exp Rheumatol* Published Online First: 2015. doi:10.1002/art.30136

36 National Institute for Health and Care Excellence. Drug treatment for rheumatoid arthritis. 2020. http://pathways.nice.org.uk/pathways/rheumatoid-arthritis

37 Kumar K, Peters S, Barton A. Rheumatoid arthritis patient perceptions on the value of predictive testing for treatments: a qualitative study. *BMC Musculoskelet Disord* Published Online First: 2016. doi:10.1186/s12891-016-1319-x

Figure Legends

Figure 1 – LCMM trajectories – Pane A reports the 2C-DAS28 trajectories of the subjects belonging to each discovered latent class, represented with locally weighted scatter plot smoothing method (LOESS) and confidence intervals. Pane B reports 2C-DAS28 individuals' trajectories in each of the identified latent classes.

Figure 2- DAS28 components in Secondary Non-Responders (pane A) and Low Activity Disease (pane B) - represented with locally weighted scatter plot smoothing method (LOESS) and confidence intervals. In Early Secondary Non-Responders both Swollen joint count and CRP follow the same pattern as detected by 2C-DAS28, in Late Secondary Non-Responders instated the non-response pattern is mainly driven by the swollen component, while the CRP pattern is mainly flat. The trajectories of Low Activity Disease group had high values of Tender Joints and Visual Analogue Scale, and the disease activity was mainly driven by these two components.

Figure 3 – Comorbidities prevalence in the LCMM subgroups. Each column indicates, for comorbidity, how the total number of subjects affected by that comorbidity are distributed in each latent class. The width of the columns indicates the prevalence of each comorbidity over the whole population.

Figure 4 – Pane A: Kaplan Mayer curves of ADAb positivity in LCMM Classes - follow-up are censored at the last registered follow-up. Pane B: Cox Regression Model ADAb

Table 1 – Baseline characteristics of the entire studied cohort.

	Total	*Missingness N (%)*
Number Of Subject	2991	
Age Mean(Sd)	62.18 (12.46)	12 (0.4)
Time From Diagnosis (Sd)	10.46 (9.75)	157 (5.24)
Height Mean(Sd)	164.49 (11.83)	589 (19.67)
Weight Mean(Sd)	79.24 (19.81)	197 (6.58)
HAQ Score Mean(Sd)	1.72 (0.60)	475 (15.86)
Sex: Male N(%)	710 (23.73)	
Ever Smoke N(%)	1868 (62.45)	432 (14.42)
2-Component Das28 Mean(Sd)	4.25 (1.19)	
Das28 Mean(Sd)	5.65 (0.96)	
Tender Mean(Sd)	14.56 (7.26)	2 (0.06)
Swollen Mean(Sd)	8.35 (5.35)	
VAS Mean(Sd)	72.01 (19.49)	68 (2.27)
CRP Mean(Sd)	20.49 (29.10)	
Treatment Combintation		
Bio+Dmards N(%)	559 (18.69)	
Bio+Dmards (Including Mtrx) N(%)	1878 (62.79)	
Only Bio N(%)	554 (18.52)	
Previously Treated With Biological	673 (22.5)	
Type Of Biological		
TNF Inhibitor N(%)	2151 (71.92)	
Adalimumab	664	
Certolizumab	373	
Etanercept	904	
Other	210	
B Cell Inhibitor N(%)	436 (14.58)	
IL-6R Inhibitor N(%)	278 (9.29)	
Other N(%)	3 (0.10)	
T-Cell Inhibitor N(%)	123 (4.11)	

Table 2 – Baseline characteristics in the 2C-DAS28 LCMM Classes

	Good Responders 1	*Good Responders 2*	*Gradual Responders 1*	*Gradual Responders 2*	*Early Secondary Non-Responders*	*Late Secondary Non-Responders*	*Low Disease Activity*	*P-Value*
Number Of Subject	395	1840	65	506	33	64	88	
Age Mean(Sd)	62.01(13.13)	62.19(12.42)	58.57(14.62)	63.01(11.63)	58.76(12.9)	62.83(13.03)	61.41(12.29)	0.09
Time From Diagnosis (Sd)	12.42(11.58)	10.45(9.64)	6.41(5.51)	9.55(9.02)	10.47(9.34)	11.33(9.07)	9.46(8.6)	<0.001
BMI Mean(Sd)	28.57(6.18)	29.39(9.52)	29.33(6.11)	29.8(8.34)	28.92(7.32)	28.76(5.23)	30.19(10.17)	0.46
HAQ Score Mean(Sd)	1.77(0.63)	1.68(0.66)	1.94(0.61)	1.8(0.61)	1.67(0.55)	1.73(0.65)	1.77(0.67)	<0.001
Sex: Male N(%)	103(26.08)	441(23.97)	16(24.62)	98(19.37)	11(33.33)	20(31.25)	21(23.86)	0.11
Ever Smoke N(%)	258(65.32)	1108(60.22)	38(58.46)	340(67.19)	30(90.91)	37(57.81)	57(64.77)	0.67
2-Component-DAS28 Mean(Sd)	5.86(0.57)	3.89(0.87)	6.31(0.74)	4.24(0.84)	5(1.16)	5.64(0.71)	1.78(0.79)	<0.001
4-Component-DAS28 Mean(Sd)	6.57(0.7)	5.43(0.87)	6.85(0.74)	5.65(0.79)	6.01(1.02)	6.34(0.81)	4.46(0.93)	<0.001
Tender Mean(Sd)	17.52(7.26)	13.83(6.98)	18.58(6.54)	14.59(7.26)	14.85(7.98)	15.98(7.53)	12.28(8.5)	<0.001
Swollen Mean(Sd)	15.18(5.21)	6.78(3.67)	17.28(5.26)	7.85(4.26)	11.24(5.79)	14.73(5.69)	1.06(1.45)	<0.001
VAS Mean(Sd)	76.66(18.5)	71.11(19.59)	76.21(17.89)	71.29(19.38)	74.31(23.92)	76.75(14.19)	66.71(21.25)	<0.001
CRP Mean(Sd)	41.9(41.79)	14.96(22.51)	55.77(45.08)	19.62(25.33)	27.87(26.22)	33.87(34.91)	6.36(10.81)	<0.001
Treatment Combination								0.911
Bio+Dmards N(%)	82(241)	341(1164)	10(41)	93(314)	4(21)	9(42)	20(55)	
Bio+Dmards(Including Mtrx) N(%)	241(61.01)	1164(63.26)	41(63.08)	314(62.06)	21(63.64)	42(65.62)	55(62.5)	
Only Bio N(%)	72(18.23)	335(18.21)	14(21.54)	99(19.57)	8(24.24)	13(20.31)	13(14.77)	
Type Of Biological								0.002

TNF Inhibitor N(%)	295(74.68)	1347(73.21)	37(56.92)	343(67.79)	21(63.64)	49(76.56)	59(67.05)	
B Cell Inhibitor N(%)	42(10.63)	241(13.1)	16(24.62)	105(20.75)	4(12.12)	8(12.5)	20(22.73)	
IL-6R Inhibitor N(%)	45(11.39)	173(9.4)	6(9.23)	37(7.31)	6(18.18)	4(6.25)	7(7.95)	
Other N(%)	0(0)	3(0.16)	0(0)	0(0)	0(0)	0(0)	0(0)	
T-Cell Inhibitor N(%)	13(3.29)	76(4.13)	6(9.23)	21(4.15)	2(6.06)	3(4.69)	2(2.27)	
Previously Treated With Biological N(%)	67(16.96)	398(21.63)	19(29.23)	136(26.88)	7(21.21)	14(21.87)	32(36.36)	<0.001
Number of Comorbidities in individuals Mean(SD)	1(1.07)	1.13(1.22)	1.08(1.08)	1.11(1.24)	1.12(1.22)	1.19(1.23)	1.31(1.23)	<0.001
Comorbidities N(%)								
High Blood Pressure	115(29.11)	534(29.02)	20(30.77)	149(29.45)	7(21.21)	23(35.94)	32(36.36)	
Angina	17(4.3)	98(5.33)	0(0)	14(2.77)	0(0)	0(0)	2(2.27)	
Heart Attack	12(3.04)	65(3.53)	0(0)	16(3.16)	1(3.03)	0(0)	3(3.41)	
Heart Failure	2(0.51)	11(0.6)	0(0)	4(0.79)	1(3.03)	0(0)	2(2.27)	
Stroke	9(2.28)	44(2.39)	3(4.62)	14(2.77)	0(0)	2(3.12)	3(3.41)	
Cancer	17(4.3)	117(6.36)	5(7.69)	27(5.34)	2(6.06)	2(3.12)	7(7.95)	
Asthma	44(11.14)	238(12.93)	11(16.92)	70(13.83)	3(9.09)	10(15.62)	21(23.86)	
Chronic Bronchitis Emphysema	23(5.82)	96(5.22)	2(3.08)	27(5.34)	8(24.24)	7(10.94)	6(6.82)	
Pneumonitis	3(0.76)	18(0.98)	0(0)	5(0.99)	0(0)	0(0)	0(0)	
Tuberculosis	5(1.27)	43(2.34)	2(3.08)	10(1.98)	1(3.03)	2(3.12)	3(3.41)	
Peptic Ulcer	12(3.04)	85(4.62)	1(1.54)	20(3.95)	3(9.09)	4(6.25)	2(2.27)	
Liver Disease	5(1.27)	41(2.23)	0(0)	10(1.98)	0(0)	0(0)	2(2.27)	
Renal Disease	7(1.77)	41(2.23)	1(1.54)	11(2.17)	0(0)	1(1.56)	1(1.14)	
Diabetes	31(7.85)	148(8.04)	2(3.08)	47(9.29)	4(12.12)	9(14.06)	7(7.95)	

Hyperthyroidism	18(4.56)	76(4.13)	4(6.15)	24(4.74)	0(0)	3(4.69)	4(4.55)
Demyelination	1(0.25)	5(0.27)	0(0)	0(0)	0(0)	0(0)	1(1.14)
Depression	73(18.48)	398(21.63)	19(29.23)	104(20.55)	5(15.15)	13(20.31)	15(17.05)
Epilepsy	1(0.25)	16(0.87)	0(0)	8(1.58)	2(6.06)	0(0)	4(4.55)

Contributorship

- Obtaining data: BRAGGS
- Study Design: Arianna Dagliati, Darren Plant, Nophar Geifman and Anne Barton
- Programming and analysis: Arianna Dagliati, Darren Plant, Beatrice Amico
- Manuscript preparation and revision: Arianna Dagliati, Darren Plant, Nisha Nair, Meghna Jani, Niels Peek, Ann W Morgan, John Isaacs, Anthony G Wilson, Kimme L Hyrich, Nophar Geifman and Anne Barton

Acknowledgements

We thank MATURA (a stratified medicine initiative jointly funded by MRC and Versus Arthritis, the NIHR Manchester BRC, MRC Molecular Pathology Node MMPathic, the NIHR Newcastle BRC and Versus Arthritis for support. ACB and JDI are NIHR Senior Investigators.

Ethical approval information

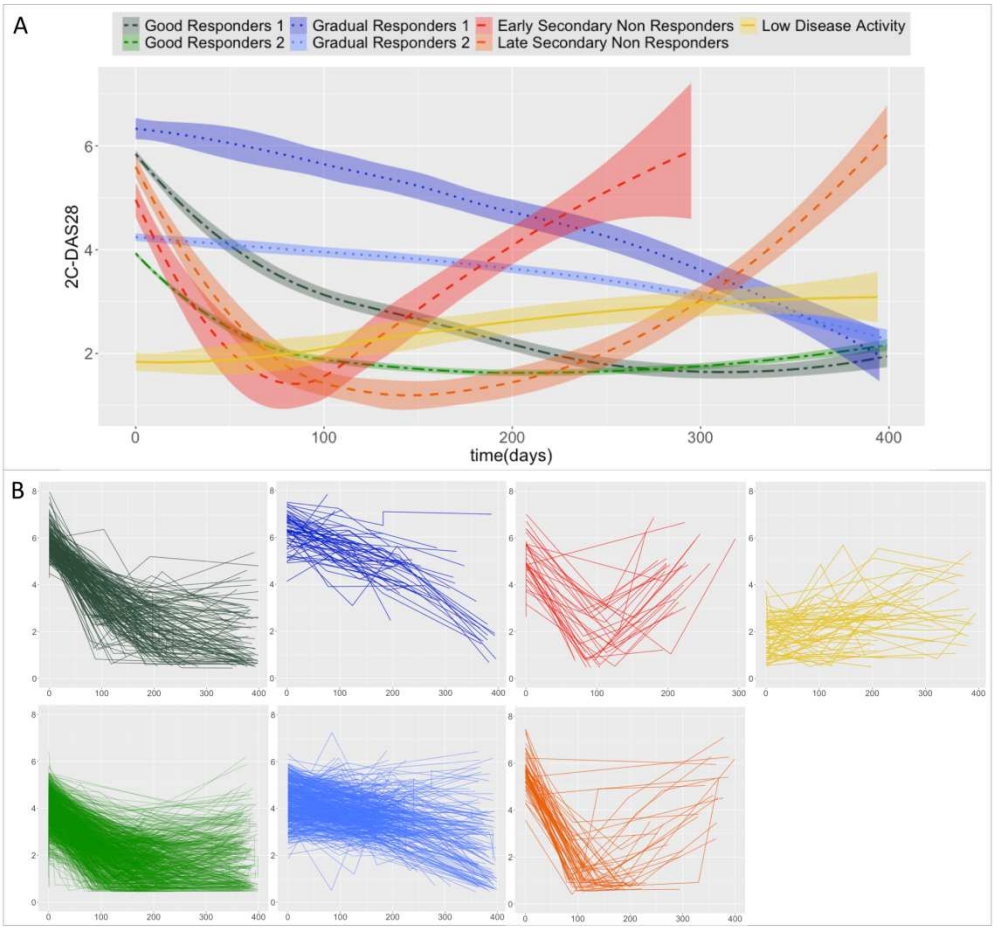
Contributing patients provide written informed consent, and the BRAGGSS study is ethically approved (COREC 04/Q1403/37).

Data sharing statement

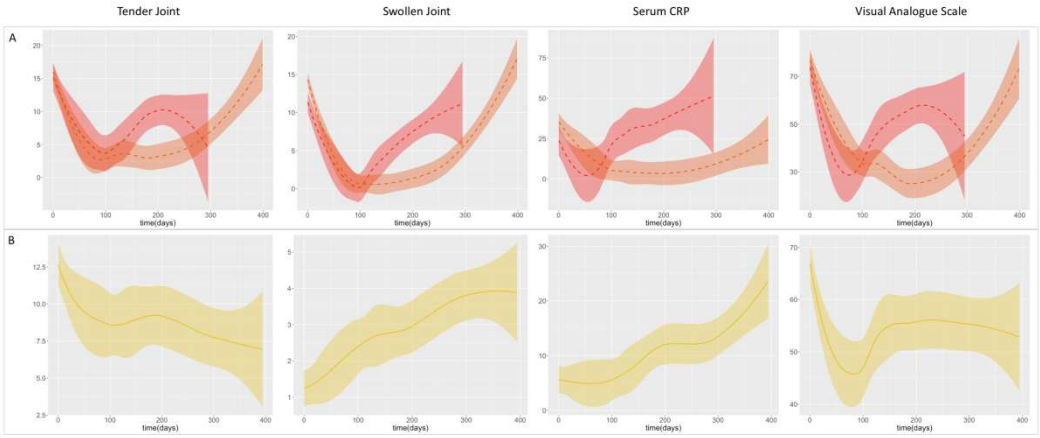
All data relevant to the study are included in the article or uploaded as supplementary information.

Patient and Public Involvement

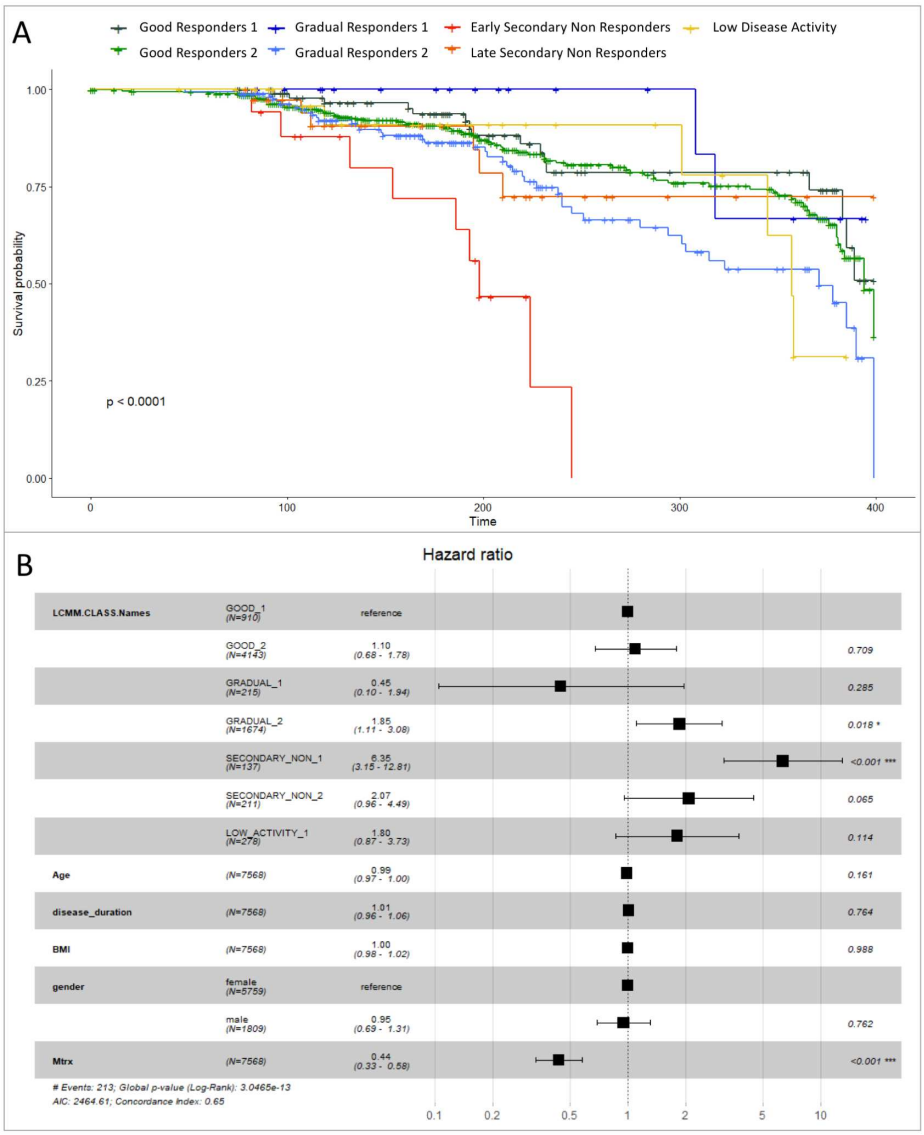
Patients were originally consulted when setting up the BRAGGSS study and continue to provide valuable feedback regarding our work on drug stratification.



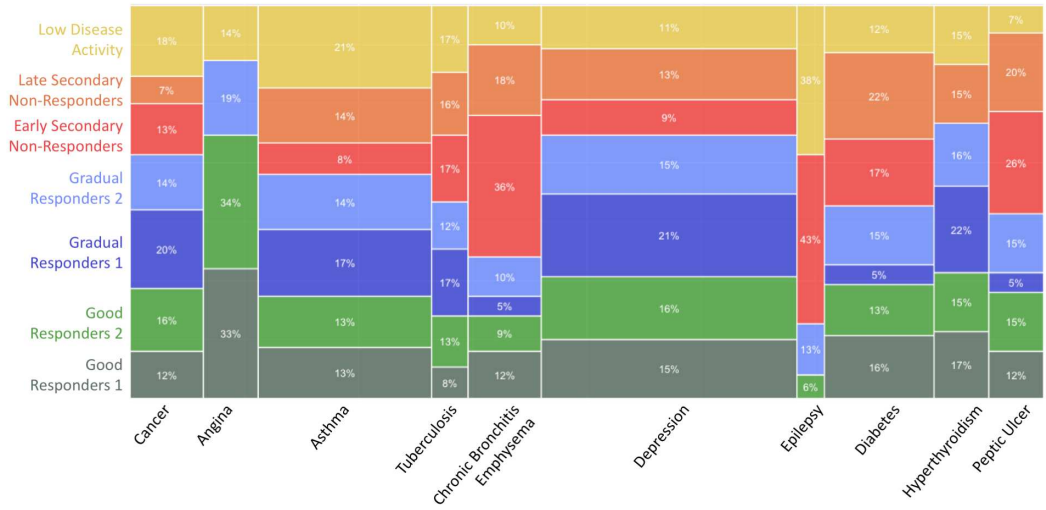
art_41379_f1.tiff



art_41379_f2.tiff



art_41379_f3.tiff



art_41379_f4.tiff