

Genome-wide Association Study of Response to Methotrexate in Early Rheumatoid Arthritis Patients

John C Taylor, MSc¹, Tim Bongartz, MD², Jonathan Massey, PhD³, Borbala Mifsud, PhD⁴, Athina Spiliopoulou, PhD⁵, Ian Scott⁶, Jianmei Wang, PhD⁷, Michael Morgan, PhD^{1,18}, Darren Plant, PhD³, Marco Colombo, PhD⁵, Sarah Twigg¹, Iain McInnes⁸, Duncan Porter⁸, Jane Freeston¹, Jackie Nam¹, Heather Cordell, PhD⁹, John Isaacs, PhD¹⁰, Jenna Strathdee, MSc¹, Donna Arnett, PhD¹¹, M.J.H. de Hair, PhD¹², Stella Aslibekyan, PhD¹³, Ronald van Vollenhoven¹⁴, S. Louis Bridges, Jr., PhD¹³, Costantino Pitzalis, PhD¹⁵, Andy Cope⁶, Suzanne Verstappen, PhD³, Paul Emery¹, Michael R Barnes, PhD¹⁵, Paul McKeigue, PhD⁵, Taisei Mushiroda, PhD¹⁶, Michiaki Kubo, PhD¹⁶, Richard Weinshilboum, MD¹⁷, Anne Barton, PhD³, Ann W Morgan, PhD¹, Jennifer H Barrett, PhD¹, on behalf of the PAMERA and MATURA Consortia†

¹ School of Medicine, University of Leeds, and NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

² Vanderbilt University, Nashville, Tennessee, USA

³ Arthritis Research UK Centre for Genetics and Genomics, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK and NIHR Manchester Musculoskeletal BRC, Central Manchester Foundation Trust, Manchester, UK

⁴ Clinical Pharmacology, William Harvey Research Institute, Queen Mary University, London, UK

⁵ Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK

⁶ King's College, London, UK

⁷ Roche Products, Welwyn Garden City, UK

⁸ Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, UK

⁹ Institute of Genetic Medicine, University of Newcastle, Newcastle, UK

¹⁰ Institute of Cellular Medicine, University of Newcastle, Newcastle, UK

¹¹ University of Kentucky College of Public Health, Lexington, KY 40536

¹² University Medical Centre Utrecht, Utrecht, The Netherlands

¹³ Comprehensive Arthritis, Musculoskeletal, Bone, and Autoimmunity Center, University of Alabama, Birmingham, USA

¹⁴ Karolinska Institute, Stockholm, Sweden

¹⁵ Queen Mary University, London, UK

¹⁶ RIKEN Center for Integrative Medical Sciences, Japan

¹⁷ Mayo Clinic, Rochester, Minnesota, USA

¹⁸ Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridge, UK

† See Supplementary Information

Corresponding Author

Professor Ann W Morgan

Leeds Institute of Rheumatic and Musculoskeletal Medicine

Wellcome Trust Brenner Building,

St. James's University Hospital

Leeds, LS9 7TF

e-mail: a.w.morgan@leeds.ac.uk

Tel no.: +44 113 343 8414

Running title: Genetics of methotrexate response in rheumatoid

Abstract (n=147; 150 allowed)

Methotrexate (MTX) monotherapy is a common first treatment for rheumatoid arthritis (RA), but many patients do not respond adequately. In order to identify genetic predictors of response, we have combined data from two consortia to carry out a genome-wide study of response to MTX in 1424 early RA patients of European ancestry. Clinical endpoints were change from baseline to 6 months after starting treatment in swollen 28-joint count, tender 28-joint count, C-reactive protein and the overall 3-component disease activity score (DAS28). No SNP reached genome-wide statistical significance for any outcome measure. The strongest evidence for association was with rs168201 in *NRG3* ($P=10^{-7}$ for change in DAS28). Some support was also seen for association with *ZMIZ1*, previously highlighted in a study of response to MTX in juvenile idiopathic arthritis. Follow-up in two smaller cohorts of RA patients did not support these findings, although these cohorts were more heterogeneous.

Introduction (3496 words excluding refs etc; 3500 allowed)

Rheumatoid arthritis (RA) is a common autoimmune inflammatory arthritis, with a significant effect on health and wellbeing. The first choice of treatment for RA patients for over a decade was monotherapy with the disease modifying anti-rheumatic drug (DMARD) methotrexate (MTX), with or without glucocorticoids¹. However, a substantial proportion of patients do not respond adequately to treatment with MTX monotherapy² at which point treatment may be escalated to combination DMARDs or biological therapy. Unfortunately, by this stage some irreversible joint damage may have occurred³. This has led to the introduction of the treat to target (T2T) approach, which advocates frequent clinical review and more rapid escalation to combination DMARD or biological therapy in order to achieve remission or low disease activity, with good evidence that these are achieved more rapidly⁴. Irrespective of the treatment protocol, MTX is the cornerstone of therapy in RA and is often continued in patients receiving biological therapies⁶.

It would be of both clinical and economic benefit to be able to target initial treatment, so that patients unlikely to respond to MTX monotherapy could be offered alternative therapies at an earlier stage. Genetic variants do not change over time and are easily measured; identifying genetic predictors of response to treatment could therefore be of great clinical utility if sufficiently predictive.

A number of investigations of genetic predictors of response to MTX in RA patients have been published, but these have largely been candidate gene studies based on genes involved in MTX metabolism, including the folate pathway, and almost all have had small sample sizes⁷⁻¹⁰. A few genetic variants have been nominally associated with response in more than one study (rs2372536 in *ATIC*, reviewed in Plant *et al.*¹¹) or in a meta-analysis (rs1051266 in *SLC19A1 (RFC1)*¹²), but results are inconsistent across studies and not close to reaching genome-wide significance levels. Senapati *et al.*¹³ carried out the first genome-wide association (GWA) analysis of response to MTX monotherapy, in 457 RA patients from North India, comparing good and poor responders; they found 2 loci reaching significance levels of less than 10^{-5} , but nothing reached genome-wide significance. Cobb *et al.*¹⁴ carried

out a GWA study of response to MTX in 759 European patients with juvenile idiopathic arthritis (JIA), identifying 14 regions reaching a significance level of less than 10^{-5} .

The clinical importance of this question motivated the formation of the international Pharmacogenomics of Methotrexate in RA (PAMERA) consortium and the UK MAximizing Therapeutic Utility in RA (MATURA) consortium (www.matura.whri.qmul.ac.uk), which has the wider remit of using blood-based biomarkers and pathobiology to inform the stratification of all stages of RA treatment. In this study, we have combined the data from these two consortia to report the first published GWA study of response to MTX monotherapy in RA patients of European ethnicity, based on a much larger sample size than previously reported studies.

Subjects and Methods

Study Population

The study was split into two phases. The first phase comprised RA cases treated with MTX monotherapy, obtained from three observational UK studies [the Yorkshire Early Arthritis Register (YEAR)¹⁵, the Rheumatoid Arthritis Medication Study (RAMS)¹⁶ and the Leeds Inflammatory Arthritis disease CONTinuum registry (IACON)¹⁷], four UK-led clinical trials [Infliximab as inDUCTION therapy in Early rheumatoid Arthritis (IDEA)¹⁸, Etanercept and Methotrexate in Patients to Induce Remission in Early arthritis trial (EMPIRE)¹⁹, Combination Anti-Rheumatic Drugs in Early Rheumatoid Arthritis (CARDERA-1)²⁰ and Effect of anakinra as combination therapy: second UK combination therapy in early rheumatoid arthritis trials (CARDERA-2)²¹], a US-led trial [Treatment of Early Aggressive Rheumatoid arthritis (TEAR)²²], a Swedish trial [SWEdish FarmacOTherapy (SWEFOT)²³] and a Dutch observational study²⁴. Although 1952 patients on MTX monotherapy were genotyped, a maximum of 1424 were available for analysis after exclusions for high steroid use, missing clinical data or quality control (Supplementary Table 1). The second phase was used for follow-up and consisted of two datasets: Phase 2a cases were obtained from the Scottish Early Rheumatoid Arthritis (SERA) study²⁵, and Phase 2b cases were from the MTX control arms of three international clinical trials programs [a 3-trial program from the Actemra versus Methotrexate double-

Blind Investigative Trial In mONotherapy (AMBITION) study²⁶, a 3-trial program involving MabThera/Rituximab²⁷⁻²⁹ and a 4-trial program involving Ocrelizumab³⁰].

For inclusion in Phase 1 of the study, individuals were required to have a consultant diagnosis of RA, a maximum of 12-month disease duration prior to starting MTX and to have started treatment with MTX monotherapy; those starting with any additional disease-modifying anti-rheumatic drugs (DMARDs), biologics or high dose oral steroid (relevant to the CARDERA studies) were excluded from analysis. The study was restricted to patients of self-reported European ancestry, further validated through SNP genotyping. Clinical measurements collected included swollen 28-joint count (SJC28), tender 28-joint count (TJC28) and C-reactive protein (CRP), or erythrocyte sedimentation rate (ESR), if CRP was not available. Measurements of these were taken at baseline (up to 6-weeks before the individual's MTX start date) and again at follow-up (6 months from the individual's MTX start date, or 3 months if either this was not available or the patient started any additional DMARD and/or biologic between 3 and 6 months). Three component disease activity scores DAS28CRP3 (calculated as $1.10[0.56(\sqrt{TJC28}) + 0.28(\sqrt{SJC28}) + 0.36(\log_e(CRP+1))] + 1.15$) and DAS28ESR3 (calculated as $1.08[0.56(\sqrt{TJC28}) + 0.28(\sqrt{SJC28}) + 0.70(\log_e(ESR))] + 0.16$) were calculated for each individual at baseline and follow-up.

For Phase 2 of the study, there was some relaxation of the entry criteria. For Phase 2a, cases were not required to have started treatment on MTX monotherapy, and additional DMARDs were permitted in combination with MTX. For Phase 2b, cases were all treated with MTX monotherapy, but were not required to have disease duration of less than 12 months.

Further information on the contributing studies is given in Supplementary Information.

Genotyping

Samples were genotyped using five different Illumina arrays (Supplementary Table 1), including (for one study) ImmunoChip, which has less than genome-wide coverage. Quality control on samples and SNPs was performed separately for each of the genotype arrays. Samples were excluded for any of the following reasons: (i) a call-rate of <98% (of the total number of SNPs on the chip); (ii) evidence of non-European origin from principal

components analysis using EIGENSTRAT³¹ after combining with HapMap and in-house European samples; (iii) sex as ascertained by genotyping not matching reported gender; (iv) evidence of first degree relationship or identity with another sample (in which case the sample with the lower call rate of the pair was excluded). SNPs were excluded for any of the following reasons: (i) Hardy–Weinberg equilibrium P -value $<10^{-6}$; (ii) call-rate $<98\%$; (iii) minor allele frequency (MAF) <0.01 . Imputation was conducted using IMPUTEv2³² with the 1,000 Genomes haplotypes Phase 3 integrated variant set as reference. Genotypes were phased within IMPUTEv2 and imputed in a 2Mb window (non-CoreExome genotyped samples) or pre-phased using SHAPEITv2-r837³³ and imputed in a 7Mb window (CoreExome genotyped samples). The MCMC options used in IMPUTEv2 for all imputation were (k=80, iter=30, burnin=10) with effective size of population set as $N_e=20000$). Following imputation, only SNPs that had an INFO score >0.8 were retained for analysis.

Ultra-low coverage whole-genome sequencing at a mean depth of 0.5x was available for the SERA participants. The raw sequencing reads were aligned to the HG19 reference genome using the Torrent 538 Mapping Alignment Program for Ion Torrent™ Data (TMAP) software. Imputation to the 1000 Genomes Phase 3 reference panel was performed using the GenImp software³⁴. Similarly to the genotype data, relatedness was evaluated based on a subset of the imputed genotypes thinned for linkage disequilibrium (LD) using PLINK³⁵, and samples were excluded on evidence of a first degree relationship with another sample.

Statistical Analyses

In Phase 1 of the study, four separate genome-wide analyses were conducted for change in outcome from baseline to follow-up (follow-up measure minus baseline measure) of DAS28, SJC28, TJC28 and $\log_e(\text{CRP}+1)$; CRP measures were log-transformed because the distribution of measures was highly positively skewed. Imputed genotypes were analyzed as expected genotype counts based on the posterior probabilities (gene dosage) using linear regression implemented in SNPTEST2³⁶, assuming an additive genetic model, with baseline measure included as a covariate. As outliers from count data may have a large influence on the fitted linear regression models for rarer SNPs, the results presented in this study are restricted to SNPs with a minor allele frequency >0.05 . Analyses were performed separately for each of the three groups of studies measured on the same genotype array (see Figure 1), and meta-

analysis was then conducted across groups using PLINK. A fixed effects model was used for the meta-analysis unless there was evidence of heterogeneity between the study arrays ($I^2 > 31\%$)³⁷, in which case a random-effects model was used. To account for potential population stratification in the OmniExpressExome samples from the USA, UK, Sweden and the Netherlands, the first two principal components were included as covariates in the regression model. The analysis of the most significant SNPs was repeated including the first 5 principal components and had little effect on the results. All outcomes were mean-centered and scaled to have a variance of 1 within each group before the linear regression. Since ESR measures were used in place of CRP measures for the CARDERA and TEAR samples, the scaling for DAS28 was carried out separately for these studies.

All regions that included a SNP with a P -value $< 1 \times 10^{-5}$ and additional SNPs within 200Kb reaching $P < 5 \times 10^{-5}$ were followed up in Phase 2. All SNPs within the region with a p -value < 0.001 were tested using the same regression models as in Phase 1. Two separate genotype arrays were used for the Phase 2b samples (Figure 1), and data from the two arrays were meta-analysed using a fixed effects model.

Follow-up of Candidate Genes and Previous Studies

Candidate genes were identified based on their putative role in MTX metabolism and mechanism of action. MTX is a structural analogue of folic acid and interferes with the folate cycle through competitive inhibition of dihydrofolate reductase, which is important for the effective metabolism of biologically active folate cofactors. These are required for the generation of methionine from homocysteine and for the synthesis of purine and pyrimidine nucleosides. Within cells, MTX is rapidly converted to γ -glutamyl polyglutamates that inhibit enzymes crucial for de novo synthesis of nucleotide precursor metabolites and ultimately inhibition of enzymes involved in adenosine metabolism. Relevant pathways, including drug transporters, were identified in the Pharmacogenomics Knowledgebase PharmGKB, using the search term “methotrexate”. The resultant pathways were “Antimetabolite Pathway” and “Methotrexate Pathway (Cancer Cell) Pharmacokinetics”. Non-redundant genes listed under these headings were included as candidates. In addition, adenosine receptor genes listed in a recent MTX pathway PharmGKB³⁸ summary were included in the analysis (Supplementary Table 2). All SNPs found within the genes were

tested in Phase 1 for each of the 4 traits.

Two previously reported genome-wide association studies of response to MTX were followed up using the results of our overall DAS28 analysis. The first was a study of 457 Indian RA patients¹³ with an outcome classified as good response (a decrease in the DAS28ESR3 score by 1.2 and the final DAS28ESR3 ≤ 3.2 for at least 6 months after MTX monotherapy) or poor response (duration of illness not exceeding 5 years and active disease (DAS28ESR3 ≥ 5.1) despite at least 3 months of therapy with MTX). The second was a study of 374 European JIA patients¹⁴ categorized as responders according to the American College of Rheumatology paediatric 30, 50 and 70 improvement criteria or as non-responders. All SNPs with $P < 0.0001$ in the discovery phase of these two studies were followed-up in our Phase 1 results for DAS28 outcome. Any such SNP which additionally reached $p < 0.01$ in Phase 1 was followed up in Phase 2.

Further details of methods used in the interpretation of results can be found in Supplementary Information.

Results

A total of 1424 RA cases passed quality control for Phase 1 (Supplementary Table 1), with the maximum number available for analysis dependent on the trait: 1392 for DAS28, 1424 for SJC28/TJC28 and 1133 for CRP. The numbers were lower for areas of the genome not covered by ImmunoChip (1244 and 1276 for DAS28 and SJC28/TJC28 respectively, with number for CRP unchanged). Demographic and clinical characteristics are presented in Table 1 and Supplementary Table 3. Starting doses of MTX ranged from 2.5 to 25mg, and highest mean baseline disease activity was observed in the patients from the CARDERA trials.

No SNP reached genome-wide statistical significance ($P < 5 \times 10^{-8}$) for any of the four outcomes. The most significant region, around the gene Neuregulin 3 (*NRG3*) on chromosome 10, reached 9.8×10^{-8} for change in DAS28 (Table 2, Figure 2, Supplementary Figures 1-2). The SNP rs168201 in this region was also the SNP with the most consistent evidence of association over the four outcomes; rs168201 is an intronic variant in *NRG3*,

having $P = 9.8 \times 10^{-8}$ for DAS28 and $P < 0.01$ for all 3 of the components SJC28, TJC28 and CRP (Figure 3).

In total, 25 regions were identified harbouring SNPs with suggestive evidence for association with MTX response ($P < 1 \times 10^{-5}$ for at least one SNP, and additional supporting SNPs within 200kb reaching $P < 5 \times 10^{-5}$). Only two regions reached this level of significance for change in CRP (around the genes *KRAS1* and *WNK2*). The analysis of SJC28 identified the most regions reaching $P < 1 \times 10^{-5}$, including regions encompassing the genes *RWDD3*, *PARK2*, *COL25A1*, *THSD7B* and *THBD*. The most significant region for TJC28 was at an intergenic region on chromosome 6 (Table 2).

There were 429 RA cases available for analysis in Phase 2a, of whom only 154 (36%) were on MTX monotherapy, the remainder being on MTX in combination with one or more of prednisolone (116), sulfasalazine (155), hydroxychloroquine sulfate (175) and leflunomide (9). For Phase 2b, 177 cases were available, with 99 (53%) of these having disease duration greater than 12 months; mean baseline disease activity was higher than for Phase 1 cases (Table 1, Supplementary Table 3). All SNPs with $P < 0.001$ in the 25 regions were followed up, but no SNP reached a level of $P < 0.001$ in either Phase 2a or 2b (Table 2, Supplementary Tables 4-7).

Details of the genes tested from the MTX metabolism candidate gene selection can be found in Supplementary Table 2. The number of SNPs tested for each gene ranged from 7 (*FPGS*) to 643 (*ADK*) with 3168 tested in total (Supplementary Figure 3). The most significant SNP was rs7996393 for both DAS28 and TJC28, found in the gene *AMPD1* ($P=0.0008$ and $p=0.0011$, respectively). For CRP, the most significant SNP was rs4148160 ($P=0.0078$) in *ABCG2*, and for SJC28 the most significant SNP was rs2236224 ($P=0.0064$) in *MTHFD1*. The genes selected included two SNPs nominally associated in more than one previous study. In *ATIC*, rs2372536 was not associated with DAS28, CRP or SJC28 ($P=0.11$, $P=0.20$ and $P=0.83$ respectively) but was nominally associated with TJC28 (beta (95% CI) = 0.06 (0.00, 0.13), $P=0.042$). In *SLC19A1*, rs1051266 was not associated with DAS28, CRP, SJC28 or TJC28 ($P=0.76$, $P=0.46$, $P=0.24$ and $P=0.42$ respectively).

We used the results from the analysis of DAS28 to follow up all SNPs with $P < 10^{-4}$ in either of two previous genome-wide association studies of response to MTX, of Indian RA patients and European JIA patients (Supplementary Table 8 and 9). No SNP from the Indian RA study showed evidence of association at $p < 0.01$ in our results when taking direction of effect into account. One region of association at the *ZMIZ1* gene in the JIA study showed some evidence of association in our results (rs703979: $P = 7.7 \times 10^{-4}$ in the JIA study, $P = 1.1 \times 10^{-4}$ for DAS28 in this study, and rs703970: $P = 8.3 \times 10^{-5}$ in the JIA study, $P = 1.8 \times 10^{-4}$ for DAS28 in this study) with corresponding directions of effect. When using Fisher's method for combining P -values across the JIA study and this study, this gives 1.4×10^{-6} for rs703979 and 2.8×10^{-7} for rs703970. These two SNPs are in almost complete LD ($r^2 = 0.98$ for the OmniExpressExome samples).

Discussion

This study is by far the largest GWA study of response to MTX in RA patients reported to date. Although no SNP showed association with response at a genome-wide level, several of the findings are of potential interest.

The SNP showing the most significant association with DAS28 (or any outcome) is rs168201, an intronic variant in *NRG3*. In an early GWA study³⁹, rs10509440 in *NRG3* was reported to show some evidence of association with RA susceptibility ($P = 6 \times 10^{-5}$), with weak supportive evidence from a neighbouring SNP, rs12358407, in a follow-up study ($P = 0.003$)⁴⁰. These findings have however not been replicated in more recent larger GWA studies. SNPs in the *SFTPD* gene close to *NRG3* have been associated with RA susceptibility at a genome-wide significance level in Asian but not European populations (rs726288)⁴¹; these SNPs are not in LD with the previously reported *NRG3* SNPs.

NRG3 encodes a cell-cell signaling protein (NRG3), which is a ligand for the ERBB4 (HER4) receptor tyrosine kinase⁴². There is some evidence linking NRG3 signaling to MTX response. Knock-down of the NRG3 receptor, ERBB4 (HER4), in osteosarcoma cell lines increased sensitivity to MTX (60% increase in apoptosis) but not other chemotherapies⁴³. In a peripheral blood monocyte study of 32 RA patients treated with anakinra and MTX, *NRG3* is

one of 7 genes in a gene signature predicting response, being up-regulated in therapeutic non-responders⁴⁴. High resolution (5kb) Hi-C data in the GM12878 lymphoblastoid cell line (see Supplementary Information) shows that rs168201 interacts most significantly with the *NRG3* promoter, underlining it as a candidate. From promoter capture Hi-C data in a GM12878 B-cell line⁴⁵, it is seen that a DNA fragment ~100kb from the LD block around rs168201 shows some evidence of interaction (CHiCAGO score = 5.41)⁴⁶ with the promoter of *MAT1A* (around 2.5Mb upstream of rs168201), encoding methionine S-adenosyltransferase (MAT), a key enzyme in the transmethylation cycle. This gene is of interest because low-dose MTX has been shown to inhibit its expression and activity⁴⁷.

Previously reported findings and candidate genes from the MTX metabolism pathway were followed up in this study, but little evidence was seen for their association with response. A number of these genes showed weak chromatin interactions in GM12878 cells with SNPs from the top 25 regions (e.g. *ADK* with rs703987, a SNP associated with change in DAS28). These interactions are listed in Supplementary Table 10.

The strongest support from candidate gene analyses was for SNPs in *ZMIZ1*, one of three genes highlighted in a GWA study of response to MTX in JIA patients¹⁴. This gene is involved in transcription factor regulation and has been associated with several autoimmune diseases, including psoriasis, Crohn's disease⁴⁸ and multiple sclerosis (MS)⁴⁹. Fewings *et al.* have recently shown that expression of *ZMIZ1* varies in response to vitamin D and to certain MS therapies and may indicate a type of immune dysregulation potentially related to therapeutic response⁵⁰.

The top 25 loci were also annotated using two approaches - high-resolution long-range chromatin interaction and correlating regional genetic scores with genetic scores for other traits (see Supplementary Information and Supplementary Table 11). A few of the candidates from the Hi-C analysis have been indicated in inflammation biology, e.g. *NINJ1* and *RWDD3*^{51, 52}. The intergenic region on chromosome 6 most strongly associated with change in TJC28 interacts with the *CCDC167* promoter in GM12878 lymphoblastoid cells, a gene which is highly expressed in the immune system. In addition rs9910936, also associated with change in TJC28, is close to several genes, but this analysis shows that the

most relevant gene is likely to be *EFTUD2*, since the SNP is an expression-quantitative-trait locus for this gene and also shows chromatin interaction. Interestingly, some of the SNPs in the *NRG3* region of association interact with the *ZMIZ1* promoter, but the interaction is much weaker than with the *NRG3* promoter, so this is not shown as a candidate in Supplementary Table 11.

Although this is by far the largest study to date, it is still small compared with current genome-wide studies of disease susceptibility. The challenge for pharmacogenetics is to assemble large cohorts of sufficiently homogeneous patients. Efforts were made in this study to minimize heterogeneity by using strict entry criteria. Outcome measures are difficult to measure consistently, and, although overall measures such as DAS28 are important clinically, the analysis was also carried out on its individual components, since some of these are more heritable⁵³ and more objective than others. For clinical interest, analyses of “responders” (any response versus no response, and good response versus no response, according to EULAR criteria⁵⁴) were also carried out (Supplementary Tables 12 and 13). Results were similar to those from the analysis of DAS28 but generally less highly significant.

This analysis suggests that no individual common variants are likely to be sufficiently predictive of response to be of clinical utility, but with larger studies it may be possible to create genetic risk scores, which, in combination with other factors, can be of use to target initial treatment. Larger sample sizes are needed, together with follow-up of the potentially interesting findings reported here.

Acknowledgements

We thank MRC/Arthritis Research UK for their joint funding of MATURA (MR/K015346/1).

The RAMS cohort was part funded by Arthritis Research UK (grant code 20385) and the NIHR Manchester Musculoskeletal BRU.

B. M. holds an MRC eMedLab Medical Bioinformatics Career Development Fellowship, funded from award MR/L016311/1. Part of this project was enabled through access to the MRC eMedLab Medical Bioinformatics infrastructure, award MR/L016311/1.

The YEAR and IACON studies were supported in part by programme grants from Arthritis Research UK (grant numbers), the National Institute for Health Research-Leeds Musculoskeletal Biomedical Research Unit (grant numbers) and Diagnostic Evaluation Co-operative (grant numbers) and British Medical Association (Doris Hillier Award). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

The IDEA study was supported by a research grant from Investigator-Initiated Studies Program of Merck Sharp & Dohme Limited. The opinions expressed in this paper are those of the authors and do not necessarily represent those of Merck Sharp & Dohme Limited.

PAMERA was supported by the US NIH Pharmacogenomics Research Network (PGRN) funded by NIGMS (U19 GM61388) and the RIKEN Center for Integrative Medical Sciences. Funded in part by the Biobank Japan Project, funded by the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary information is available at *The Pharmacogenomics Journal's* website.

References

1. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, *et al.* 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016; **68**(1): 1-26.

2. van Vollenhoven RF. Treatment of rheumatoid arthritis: state of the art 2009. *Nat Rev Rheumatol* 2009; **5**(10): 531-541.
3. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, *et al.* Treating rheumatoid arthritis to target: recommendations of an international task force. *Annals of the rheumatic diseases* 2010; **69**(4): 631-637.
4. Schoels M, Knevel R, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas DT, *et al.* Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. *Annals of the rheumatic diseases* 2010; **69**(4): 638-643.
5. Stoffer MA, Schoels MM, Smolen JS, Aletaha D, Breedveld FC, Burmester G, *et al.* Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. *Annals of the rheumatic diseases* 2016; **75**(1): 16-22.
6. Hyrich KL, Watson KD, Silman AJ, Symmons DP, British Society for Rheumatology Biologics R. Predictors of response to anti-TNF-alpha therapy among patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. *Rheumatology (Oxford)* 2006; **45**(12): 1558-1565.
7. Dervieux T, Greenstein N, Kremer J. Pharmacogenomic and metabolic biomarkers in the folate pathway and their association with methotrexate effects during dosage escalation in rheumatoid arthritis. *Arthritis Rheum* 2006; **54**(10): 3095-3103.
8. Lee YC, Cui J, Costenbader KH, Shadick NA, Weinblatt ME, Karlson EW. Investigation of candidate polymorphisms and disease activity in rheumatoid arthritis patients on methotrexate. *Rheumatology (Oxford)* 2009; **48**(6): 613-617.
9. Jekic B, Lukovic L, Bunjevacki V, Milic V, Novakovic I, Damnjanovic T, *et al.* Association of the TYMS 3G/3G genotype with poor response and GGH 354GG genotype with the bone marrow toxicity of the methotrexate in RA patients. *Eur J Clin Pharmacol* 2013; **69**(3): 377-383.
10. Lima A, Monteiro J, Bernardes M, Sousa H, Azevedo R, Seabra V, *et al.* Prediction of methotrexate clinical response in Portuguese rheumatoid arthritis patients: implication of MTHFR rs1801133 and ATIC rs4673993 polymorphisms. *Biomed Res Int* 2014; **2014**: 368681.
11. Plant D, Wilson AG, Barton A. Genetic and epigenetic predictors of responsiveness to treatment in RA. *Nat Rev Rheumatol* 2014; **10**(6): 329-337.

12. Kung TN, Dennis J, Ma Y, Xie G, Bykerk V, Pope J, *et al.* RFC1 80G>A is a genetic determinant of methotrexate efficacy in rheumatoid arthritis: a human genome epidemiologic review and meta-analysis of observational studies. *Arthritis Rheumatol* 2014; **66**(5): 1111-1120.
13. Senapati S, Singh S, Das M, Kumar A, Gupta R, Kumar U, *et al.* Genome-wide analysis of methotrexate pharmacogenomics in rheumatoid arthritis shows multiple novel risk variants and leads for TYMS regulation. *Pharmacogenet Genomics* 2014; **24**(4): 211-219.
14. Cobb J, Cule E, Moncrieffe H, Hinks A, Ursu S, Patrick F, *et al.* Genome-wide data reveal novel genes for methotrexate response in a large cohort of juvenile idiopathic arthritis cases. *Pharmacogenomics J* 2014; **14**(4): 356-364.
15. Conaghan PG, Hensor EM, Keenan AM, Morgan AW, Emery P, Consortium Y. Persistently moderate DAS-28 is not benign: loss of function occurs in early RA despite step-up DMARD therapy. *Rheumatology (Oxford)* 2010; **49**(10): 1894-1899.
16. Hope HF, Anderson J, Barton A, Cordingley L, Hyrich K, Verstappen SMM. Anxiety and Depression Are Associated with Non-Adherence to Methotrexate in the First 6 Months of Treatment in a Rheumatoid Arthritis Population. *Rheumatology* 2015; **54**: 61-62.
17. Dumitru RB, Horton S, Hodgson R, Wakefield RJ, Hensor EM, Emery P, *et al.* A prospective, single-centre, randomised study evaluating the clinical, imaging and immunological depth of remission achieved by very early versus delayed Etanercept in patients with Rheumatoid Arthritis (VEDERA). *BMC Musculoskelet Disord* 2016; **17**: 61.
18. Nam JL, Villeneuve E, Hensor EM, Conaghan PG, Keen HI, Buch MH, *et al.* Remission induction comparing infliximab and high-dose intravenous steroid, followed by treat-to-target: a double-blind, randomised, controlled trial in new-onset, treatment-naive, rheumatoid arthritis (the IDEA study). *Annals of the rheumatic diseases* 2014; **73**(1): 75-85.
19. Nam JL, Villeneuve E, Hensor EM, Wakefield RJ, Conaghan PG, Green MJ, *et al.* A randomised controlled trial of etanercept and methotrexate to induce remission in early inflammatory arthritis: the EMPIRE trial. *Annals of the rheumatic diseases* 2014; **73**(6): 1027-1036.

20. Choy EH, Smith CM, Farewell V, Walker D, Hassell A, Chau L, *et al.* Factorial randomised controlled trial of glucocorticoids and combination disease modifying drugs in early rheumatoid arthritis. *Annals of the rheumatic diseases* 2008; **67**(5): 656-663.
21. Scott IC, Ibrahim F, Simpson G, Kowalczyk A, White-Alao B, Hassell A, *et al.* A randomised trial evaluating anakinra in early active rheumatoid arthritis. *Clin Exp Rheumatol* 2016; **34**(1): 88-93.
22. Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St Clair EW, *et al.* A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. *Arthritis Rheum* 2012; **64**(9): 2824-2835.
23. van Vollenhoven RF, Ernestam S, Geborek P, Petersson IF, Coster L, Waltbrand E, *et al.* Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (Swefot trial): 1-year results of a randomised trial. *Lancet* 2009; **374**(9688): 459-466.
24. de Hair MJ, Lehmann KA, van de Sande MG, Maijer KI, Gerlag DM, Tak PP. The clinical picture of rheumatoid arthritis according to the 2010 American College of Rheumatology/European League Against Rheumatism criteria: is this still the same disease? *Arthritis Rheum* 2012; **64**(2): 389-393.
25. Kronisch C, McLernon DJ, Dale J, Paterson C, Ralston SH, Reid DM, *et al.* Brief Report: Predicting Functional Disability: One-Year Results From the Scottish Early Rheumatoid Arthritis Inception Cohort. *Arthritis Rheumatol* 2016; **68**(7): 1596-1602.
26. Jones G, Sebba A, Gu J, Lowenstein MB, Calvo A, Gomez-Reino JJ, *et al.* Comparison of tocilizumab monotherapy versus methotrexate monotherapy in patients with moderate to severe rheumatoid arthritis: the AMBITION study. *Annals of the rheumatic diseases* 2010; **69**(1): 88-96.
27. Emery P, Deodhar A, Rigby WF, Isaacs JD, Combe B, Racewicz AJ, *et al.* Efficacy and safety of different doses and retreatment of rituximab: a randomised, placebo-controlled trial in patients who are biological naive with active rheumatoid arthritis and an inadequate response to methotrexate (Study Evaluating Rituximab's Efficacy in MTX iNadequate rEsponders (SERENE)). *Annals of the rheumatic diseases* 2010; **69**(9): 1629-1635.
28. Rubbert-Roth A, Tak PP, Zerbini C, Tremblay JL, Carreno L, Armstrong G, *et al.* Efficacy and safety of various repeat treatment dosing regimens of rituximab in

patients with active rheumatoid arthritis: results of a Phase III randomized study (MIRROR). *Rheumatology (Oxford)* 2010; **49**(9): 1683-1693.

29. Tak PP, Rigby WF, Rubbert-Roth A, Peterfy CG, van Vollenhoven RF, Stohl W, *et al.* Inhibition of joint damage and improved clinical outcomes with rituximab plus methotrexate in early active rheumatoid arthritis: the IMAGE trial. *Annals of the rheumatic diseases* 2011; **70**(1): 39-46.
30. Emery P, Rigby W, Tak PP, Dorner T, Olech E, Martin C, *et al.* Safety with ocrelizumab in rheumatoid arthritis: results from the ocrelizumab phase III program. *PloS one* 2014; **9**(2): e87379.
31. Reich D, Price AL, Patterson N. Principal component analysis of genetic data. *Nature genetics* 2008; **40**(5): 491-492.
32. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics* 2009; **5**(6): e1000529.
33. Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype estimation using sequencing reads. *American journal of human genetics* 2013; **93**(4): 687-696.
34. Spiliopoulou A, Colombo M, Orchard P, Agakov F, McKeigue P. GenImp: Fast Imputation to Large Reference Panels Using Genotype Likelihoods from Ultra-Low Coverage Sequencing. *Genetics* 2017.
35. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 2007; **81**(3): 559-575.
36. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature genetics* 2007; **39**(7): 906-913.
37. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**(11): 1539-1558.
38. Mikkelsen TS, Thorn CF, Yang JJ, Ulrich CM, French D, Zaza G, *et al.* PharmGKB summary: methotrexate pathway. *Pharmacogenet Genomics* 2011; **21**(10): 679-686.

39. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, *et al.* Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nature genetics* 2007; **39**(12): 1477-1482.
40. Julia A, Ballina J, Canete JD, Balsa A, Tornero-Molina J, Naranjo A, *et al.* Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. *Arthritis Rheum* 2008; **58**(8): 2275-2286.
41. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014; **506**(7488): 376-381.
42. Kataoka H. EGFR ligands and their signaling scissors, ADAMs, as new molecular targets for anticancer treatments. *J Dermatol Sci* 2009; **56**(3): 148-153.
43. Rivera-Valentin RK, Hua YQ, Zhang Y, Lala N, Yang YW, Hughes DPM. Her-4 mediates anoikis resistance, chemoresistance, and metastatic potential in osteosarcoma. *Cancer research* 2014; **74**(20).
44. Bansard C, Lequerre T, Derambure C, Vittecoq O, Hiron M, Daragon A, *et al.* Gene profiling predicts rheumatoid arthritis responsiveness to IL-1Ra (anakinra). *Rheumatology (Oxford)* 2011; **50**(2): 283-292.
45. Martin P, McGovern A, Orozco G, Duffus K, Yarwood A, Schoenfelder S, *et al.* Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat Commun* 2015; **6**: 10069.
46. Cairns J, Freire-Pritchett P, Wingett SW, Varnai C, Dimond A, Plagnol V, *et al.* CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biol* 2016; **17**(1): 127.
47. Wang YC, Chiang EP. Low-dose methotrexate inhibits methionine S-adenosyltransferase in vitro and in vivo. *Mol Med* 2012; **18**: 423-432.
48. Ellinghaus D, Ellinghaus E, Nair RP, Stuart PE, Esko T, Metspalu A, *et al.* Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *American journal of human genetics* 2012; **90**(4): 636-647.

49. International Multiple Sclerosis Genetics C, Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kempainen A, *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics* 2013; **45**(11): 1353-1360.
50. Fewings NL, Gatt PN, McKay FC, Parnell GP, Schibeci SD, Edwards J, *et al.* The autoimmune risk gene ZMIZ1 is a vitamin D responsive marker of a molecular phenotype of multiple sclerosis. *J Autoimmun* 2017; **78**: 57-69.
51. Jennewein C, Zacharowski K. Abstracts of Sepsis 2012. Paris, France. November 14-16, 2012. *Crit Care* 2012; **16 Suppl 3**: P22.
52. Arciuch VGA, Tedesco L, Fuertes M, Arzt E. Role of RSUME in inflammation and cancer. *Febs Lett* 2015; **589**(22): 3330-3335.
53. Umicevic Mirkov M, Janss L, Vermeulen SH, van de Laar MA, van Riel PL, Guchelaar HJ, *et al.* Estimation of heritability of different outcomes for genetic studies of TNFi response in patients with rheumatoid arthritis. *Annals of the rheumatic diseases* 2015; **74**(12): 2183-2187.
54. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996; **39**(1): 34-40.

Figure legends

Figure 1. Study design. Study names are given in italics. (*) indicates ESR was used in the DAS28 calculation for that study instead of CRP.

Figure 2. Results for the SNP (rs168201) reaching the highest level of significance from the linear regression of change in DAS28 in phase 1 of the study. (a) $-\log_{10}$ P -values from Phase 1 for the SNPs surrounding rs168201 on chromosome 10. (b) Forest plot of estimated per allele betas for the effect of rs168201 on change in DAS28 in number of standard deviations by study phase.

Figure 3. Results in all traits for the lead SNP of the 25 most highly associated regions identified in Phase 1 in DAS28, CRP, SJC28 or TJC.



