Association of response to TNF inhibitors in rheumatoid arthritis with quantitative trait loci for CD40 and CD39

Citation for published version:
Spiliopoulou, A, Colombo, M, Plant, D, Nair, N, Cui, J, Coenen, MJ, Ikari, K, Yamanaka, H, Saevarsdottir, S, Padyukov, L, Bridges, SL, Kimberly, RP, Okada, Y, van Riel, PLC, Wolbink, G, van der Horst-Bruinsma, IE, de Vries, N, Tak, PP, Ohmura, K, Canhão, H, Guchelaar, H-J, Huizinga, TW, Criswell, LA, Raychaudhuri, S, Weinblatt, ME, Wilson, AG, Mariette, X, Isaacs, JD, Morgan, AW, Pitzalis, C, Barton, A & McKeigue, P 2019, 'Association of response to TNF inhibitors in rheumatoid arthritis with quantitative trait loci for CD40 and CD39', Annals of the Rheumatic Diseases. https://doi.org/10.1136/annrheumdis-2018-214877

Digital Object Identifier (DOI):
10.1136/annrheumdis-2018-214877

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Annals of the Rheumatic Diseases

Publisher Rights Statement:
This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
TRANSLATIONAL SCIENCE

Association of response to TNF inhibitors in rheumatoid arthritis with quantitative trait loci for CD40 and CD39

Athina Spiliopoulou,1,2 Marco Colombo,1 Darren Plant,3,4 Nisha Nair,3 Jing Cui,5 Marieke JH Coenen,6 Katsunori Ikari,7,8 Hisashi Yamanaka,9 Saedis Saevarsdottir,10,11 Leonid Padyukov,10 S Louis Bridges Jr,12 Robert P Kimberly,12 Yukinori Okada,13,14 Piet L CM van Riel,6 Gertjan Wolbink,15 Irene E van der Horst-Bruinsma,16 Nick de Vries,17 Paul P Tak,17 Koichiro Ohmura,18 Helena Canhão,19 Henk-Jan Guchelaar,20 Tom WJ Huizinga,21 Lindsey A Criswell,22 Soumya Raychaudhuri,5,23 Michael E Weinblatt,5 Anthony G Wilson,24 Xavier Mariette,25 John D Isaacs,26,27 Ann W Morgan,28,29 Costantino Pitzalis,30 Anne Barton,3,4 Paul McKeigue1

ABSTRACT

Objectives We sought to investigate whether genetic effects on response to TNF inhibitors (TNFi) in rheumatoid arthritis (RA) could be localised by considering known genetic susceptibility loci for relevant traits and to evaluate the usefulness of these genetic loci for stratifying drug response.

Methods We studied the relation of TNFi response, quantified by change in swollen joint counts (ΔSJC) and erythrocyte sedimentation rate (ΔESR) with locus-specific scores constructed from genome-wide association study summary statistics in 2938 genotyped individuals: 37 scores for RA; scores for 19 immune cell traits; scores for expression or methylation of 93 genes with previously reported associations between transcript level and drug response. Multivariate associations were evaluated in penalised regression models by cross-validation.

Results We detected a statistically significant association between ΔSJC and the RA score at the CD40 locus (p=0.0004) and an inverse association between ΔSJC and the score for expression of CD39 on CD4 T cells (p=0.00005). A previously reported association between CD39 expression on regulatory T cells and response to methotrexate was in the opposite direction. In stratified analysis by concomitant methotrexate treatment, the inverse association was stronger in the combination therapy group and dissipated in the TNFi monotherapy group. Overall, ability to predict TNFi response from genotypic scores was limited, with models explaining less than 1% of phenotypic variance.

Conclusions The association with the CD39 trait is difficult to interpret because patients with RA are often prescribed TNFi after failing to respond to methotrexate. The CD39 and CD40 pathways could be relevant for targeting drug therapy.

INTRODUCTION

Biologic therapies have transformed the outlook for rheumatoid arthritis (RA). However, for the most commonly used class of agent, tumour necrosis factor inhibitors (TNFi), there is substantial variability in response to treatment among patients with RA.1 This has spurred efforts to discover predictors of response and more generally to understand how to subtype this heterogeneous disease to predict which therapies will work.2,3 Genome-wide association studies (GWAS) of response to TNFi have shown that common single
nucleotide polymorphisms (SNPs) explain an estimated 40% and 50% of the variance of change in swollen joint counts (SJC) and erythrocyte sedimentation rate (ESR), respectively; however, no strong associations with individual SNPs have been detected.4 Thus, as with many complex phenotypes, the genetic architecture of response to TNFi is likely to be polygenic with many small genetic effects.5 In this situation, the sample size required to learn a predictive model is very large—up to 10 cases per variable—and it may not be feasible to assemble such large sample sizes for studying response to a single drug or drug class.

It has been suggested that improving prediction of complex clinical outcomes may be possible by incorporating information about the genetics of relevant traits in the prediction model.7 8 One such approach is to use publicly available summary GWAS results of relevant traits to compute genotypic scores, which can then be used as variables (‘features’) from which to build predictive models. By harnessing the genetic profiles of intermediate traits, these scores aggregate the effects of individual SNPs into larger regional or whole-genome effects. Relevant traits can include diseases, biomarkers and gene transcription levels. For polygenic traits such as RA, for which multiple genetic susceptibility loci have been identified, we can construct locus-specific scores allowing us to examine the extent to which drug response is related to genetic heterogeneity of the disease.

In the current study, we incorporated available genetic information on susceptibility to RA,9 immune cell traits from a publicly available bioresource10 and expression or methylation of genes implicated in response to TNFi treatment in RA.11 The genotypic scores associated with these intermediate traits were then tested for association with response to TNFi; by reducing the number of hypotheses being explored, the thresholds for claiming statistical significance are relaxed, which could help identify useful predictors.

MATERIALS AND METHODS

Cohorts

We used a sample of 2938 individuals of European ancestry for whom complete clinical and GWAS data were available. This sample comprised individuals from a pre-existing international collaboration formed to study the genetics of response to TNFi agents12 and individuals recruited to the Biologics in Rheumatoid Arthritis Genetics and Genomics Study Syndicate (BRAGGSS) after 2013.9

Table 1 shows sample sizes, phenotypes and clinical variables for each of the data collections used in this study. All participants provided informed consent, and institutional review board/ethics approvals were in place as described in Cui et al12 and Massey et al.4

Definition of response to TNFi treatment

In RA, response to treatment is quantified by change in the Disease Activity Score (DAS), which depends on four measurements: ESR, SJC, tender joint count (TJC) and patient global health assessment rated on a visual analogue scale (GHVAS). Previous work has shown that only the SJC and ESR measurements have evidence of non-zero heritability7 and correlate significantly with synovitis quantified by ultrasound or MRI.11 14 Since TNFi were developed to control synovitis, we used the two objective components of the DAS (ESR and SJC) as primary outcomes for evaluating genetic effects and the two subjective components (TJC and GHVAS) and the composite score (DAS28-ESR4) as secondary outcomes. For each outcome, a baseline measurement was taken before initiation of TNFi treatment, and a follow-up measurement was taken between 3 and 6 months after initiation of TNFi treatment. The measurements for each component were transformed in accordance to the DAS28-ESR4 formula (see online supplementary methods). Response was modelled as the difference between the baseline and the follow-up measurement.

Genotypic risk scores

We used the GENOSCORES platform (https://pm2.phs.ed.ac.uk/genoscores/) to compute genotypic risk scores for the intermediate traits. GENOSCORES is a database of published SNP to trait associations from a large number of well-powered GWAS, including GWAS of disease traits, biomarkers, gene expression and methylation. The database is accompanied by a software package implemented in R that can be used to compute genotypic risk scores and run downstream statistical analyses in cohorts with SNP data.

We queried the GENOSCORES database for genetic associations with RA risk,9 149 heritable immune cell traits reported by Roederer et al,10 and whole-blood expression and methylation for 93 genes reported in a recent meta-analysis15 as differentially expressed before treatment between responder and non-responder patients with RA treated with TNFi. GWAS summary statistics reported by Westra et al16 and Gusev et al17 were used for expression quantitative trait loci (eQTLs). GWAS summary statistics reported by Gaunt et al15 were used for methylation quantitative trait loci (mQTLs).

GWAS summary statistics for each intermediate trait were filtered at p value<10−5. SNPs were then split into trait-associated regions, with regions defined as genomic loci containing at least one SNP with p value<10−7. Only 19 of the immune cell traits had a corresponding trait-associated region. SNPs not assigned to a region were discarded. For each trait-associated region, a genotypic score was computed as a sum of SNP genotypes, g, weighted by the effect size estimates, β (log OR for binary traits, regression slope for quantitative traits) and adjusted for linkage disequilibrium. The regional score, sR, for an individual i, trait t and region r was computed as: sR = βR t g R t r

, where R denotes the SNP–SNP correlation matrix in genomic region r.

Additional details about the GENOSCORES platform, the score computation and the specific regional scores used in this study are given in online supplementary materials (see online supplementary tables S1–S5, online supplementary figure S1).

Predictive modelling

To evaluate genetic prediction of response to TNFi, we compared a model with clinical covariates only to a model with clinical covariates and genotypic scores for each type of intermediate trait. To avoid numerical instabilities, we removed highly correlated scores prior to fitting a model (see online supplementary methods). The number of filtered regional scores for each type of intermediate trait is shown in table 2.

We expected that only a subset of genotypic scores would be relevant for prediction of response to TNFi, and thus used a hierarchical shrinkage prior for the score coefficients. We implemented the prediction models in STAN18 using a horseshoe prior distribution and performed inference with Markov chain Monte Carlo sampling.19 20 To rank the importance of genotypic scores in a model, we applied projection predictive variable selection, an approach that projects posterior draws from the high-dimensional model to lower dimensional subspaces.21
| BRAGGSS | DREAM | EIRA | ReAct | WTCCC | Other* | Total | Sample size |
|---------|--------|------|-------|-------|--------|-------|-------------|
| 954     | 764    | 283  | 258   | 556   | 123    | 2938  |             |
| 76      | 68     | 74   | 77    | 77    | 82     | 74    |             |
| 85      | 74     | 74   | 50    | 73    | 97     | 76    |             |
| 29      | 1255   | 34   | 0     | 0     | 0      | 0     |             |
| 0       | 34     | 293  | 66    | 97    | 246    |       |             |
| 19      | 721    | 17   | 0     | 0     | 0      | 0     |             |
| 0       | 17     | 194  | 138   | 139   | 246    |       |             |
| 75      | 792    |      |       |       |        |       |             |
| 6.7 (0.9) | 5.6 (1.1) | 6.0 (1.2) |       |       |       |       |             |
| 45.0 (28.8) | 30.2 (21.5) | 34.6 (25.4) |       |       |       |       |             |
| 11.8 (6.4) | 10.5 (6.3) | 10.4 (6.0) |       |       |       |       |             |
| 17.0 (7.3) | 10.7 (6.3) | 13.2 (7.8) |       |       |       |       |             |
| 72.6 (18.1) | 59.3 (23.9) | 66.1 (21.4) |       |       |       |       |             |
| 3.7 (1.4) | 4.2 (1.5) | 3.9 (1.6) | 3.8 (1.5) |       |       |       |             |
| 18.6 (17.0) | 27.6 (24.9) | 22.2 (19.7) | 21.8 (20.7) |       |       |       |             |
| 3.6 (3.6) | 4.0 (4.7) | 5.0 (4.9) | 3.8 (4.3) |       |       |       |             |
| 4.7 (5.5) | 6.3 (6.6) | 5.1 (6.1) | 4.7 (5.8) |       |       |       |             |
| 31.8 (25.9) | 37.1 (25.0) | 31.8 (26.6) | 35.5 (24.5) |       |       |       |             |

Table 1: Sample information per cohort.

- **BRAGGSS**: Biologics in Rheumatoid Arthritis Genetics and Genomics Study Syndicate
- **DREAM**: Dutch Rheumatoid Arthritis Monitoring Registry
- **EIRA**: Swedish Epidemiological Investigation of Rheumatoid Arthritis
- **ReAct**: French Research in Active Rheumatoid Arthritis
- **WTCCC**: Wellcome Trust Case Control Consortium

Sample size includes the total number of participants across all cohorts.

Sex, female %

Concomitant non-biologic DMARD %

Concomitant non-biologic DMARD includes TNF inhibitors and Janus kinase inhibitors.

Baseline disease activity, mean (SD)

ESR, Erythrocyte sedimentation rate

GHVAS, Global health assessment rated on a visual analogue scale

DAS28-ESR4, Disease Activity Score 28 joints, C-reactive protein

DAS28-ESR4, Disease Activity Score 28 joints, erythrocyte sedimentation rate

* ‘Other’ displays aggregate sample characteristics for collections with sample size <100.
Rheumatoid arthritis

Table 2 Prediction of response to TNFi using penalised regional genotypic scores for different types of intermediate traits

| Intermediate trait type | No of regional scores | No of filtered scores | Prediction of ∆SJC (%) | Prediction of ∆ESR (%) |
|-------------------------|-----------------------|-----------------------|------------------------|------------------------|
| Rheumatoid arthritis    | 37                    | 37                    | 5.3 (0.26)             | −1.6 (0)               |
| Immune cell traits      | 508                   | 470                   | −0.7 (0)               | 2.9 (0.17)             |
| eQTLs                   | 94                    | 87                    | 3.4 (0.16)             | 2.9 (0.17)             |
| mQTLs and eQTLs         | 268                   | 228                   | 2.5 (0.11)             | 1.6 (0.09)             |

Prediction performance is quantified by the difference in test log-likelihood (in nats) between a model with clinical covariates and genotypic scores and a model with clinical covariates only and by the per cent of phenotypic variance explained (in parenthesis). Results from 10-fold cross-validation.

We used a statistical model with the following clinical covariates: measurements for the four DAS components before initiation of TNFi treatment, gender, whether the patient was concomitantly treated with any non-biologic disease-modifying antirheumatic drugs (DMARDs), cohort (which is also a proxy for country), genotyping array and the 10 first principal components computed from the genotypic data of the full data set. We used individuals with complete measurements in the statistical analyses of each TNFi response outcome (see online supplementary table S6).

Evaluation of prediction

We used two measures to quantify improvement in prediction: the difference in log-likelihood between a model with clinical covariates and genotypic scores and a model with clinical covariates only (measured in natural log units (nats)); and the per cent of residual variance explained by the genotypic scores. Both measures were computed on the testing data from 10-fold cross-validation on the full data set.

Figure 1 Contribution of top 10 RA regional scores to prediction of response to TNFi phenotypes, starting from a model containing only clinical covariates. The curve gradually converges to one with the addition of all remaining scores. ESR, erythrocyte sedimentation rate; RA, rheumatoid arthritis; SJC, swollen joint count; TNFi, tumour necrosis factor inhibitors.

For readers who prefer a frequentist interpretation, the asymptotic equivalence of model choice by cross-validation and Akaike’s information criterion (AIC)2 implies that a p value of 0.01 for comparison of nested models is equivalent to a difference in test log-likelihood of 2.3 nats (likelihood ratio of 10) for models differing by one extra parameter. The large sample size of this study means that small robust increments in predictive performance can be detected.

Univariate associations

For models with a test log-likelihood difference of at least two nats, we further examined the univariate associations between genotypic scores and the TNFi response outcomes. We used the full data set to test the univariate association and included the same clinical covariates as in the multivariate prediction, in addition to a score.

For genotypic scores significantly associated with TNFi response at the Bonferroni-corrected p value threshold, we compared the estimated effects among groups receiving different TNFi agents. We considered etanercept, adalimumab and infliximab, where a large enough sample size was available. Additionally, we tested if the associations held when we adjusted for additional covariates: anticitrullinated protein antibodies (ACPAs) status and smoking status. These covariates have been reported to influence TNFi response but were only available for a third of the samples, and thus were not included in the full models.

A diagram of the statistical analysis pipeline is given in online supplementary figure S2.

RESULTS

In the following sections, we focus on the results for the primary TNFi response outcomes. The results for secondary outcomes are discussed in the online supplementary results (see online supplementary table S7).

RA genotypic scores

Prediction of response to TNFi as quantified by ∆SJC improved by including the regional genotypic scores for RA risk in a penalised regression model (table 2). The test log-likelihood increased by 5.3 nats, suggesting that some of the genetic drivers for RA are also influencing response to TNFi; however, the absolute improvement in prediction was small, with less than 1% of phenotypic variance being explained by the RA genotypic scores.

The regional score at the CD40 locus had the highest explanatory power for both response phenotypes (figure 1). The direction of the effect was consistent for response as measured by both phenotypes, with higher RA load at the CD40 locus being associated with better TNFi response. The univariate association of the RA score at the CD40 locus with ∆SJC passed the p value threshold corrected for the number of RA scores and two response phenotypes (table 3).

The regional score for RA at the CD40 locus was correlated with a cis-acting eQTL score for CD40 expression in whole blood (correlation=0.65) and a cis-acting mQTL score for methylation of CD40 in whole blood (correlation=−0.70). The strongest association between response to TNFi and genotypic scores at the CD40 locus was with the score for RA (table 3). The estimated effect did not change when we stratified by TNFi agent and when we adjusted for ACPA and smoking status (see online supplementary tables S8 and S9).

Genotypic scores for immune cell traits

Prediction of response to TNFi as measured by ∆ESR improved by adding penalised genotypic scores for immune cell traits in
Information on whether a patient was receiving a concomitant non-biologic DMARD was available for all samples, while information on whether a patient was specifically receiving concomitant MTX treatment was available for a subset of patients from the BRAGGSS cohort. Table 4 shows the effect of the genotypic score for expression of ‘CD39 on CD4 T cells’ on ΔSJC for patients receiving TNFi treatment stratified by concomitant treatment with either any non-biologic DMARD (top) or specifically with MTX (bottom). The effect became stronger in the groups receiving concomitant treatment and attenuated in the group receiving TNFi monotherapy; the CIs among all groups overlapped. Similarly, we did not detect statistically significant differences when we stratified by TNFi agent and when we adjusted for ACPA status (see online supplementary tables S8 and S9).

### eQTL and mQTL scores
Prediction of response to TNFi as quantified by both ΔSJC and ΔESR improved by adding eQTL and mQTL scores of implicated genes in a penalised regression model (table 2). Of the 93 genes reported in Kim et al.23 as differentially expressed between responders and non-responders to TNFi in RA, 54 genes had at least one eQTL, 54 genes had at least one mQTL and 36 had both. The test log-likelihood increased by 3.4 and 2.9 nats for ΔSJC and ΔESR, respectively, by adding the eQTL scores in the model. We did not see a further improvement by adding genotypic scores for mQTLs.

### DISCUSSION
In the largest international study of TNFi response to date, we have shown how using methods that leverage information from relevant intermediate traits can identify predictors of TNFi response. In a recent crowdsourced effort to use machine learning to construct a predictor of response to TNFi in RA, including SNP genotypes did not improve prediction beyond that obtained with clinical covariates alone.24 Genotypic prediction of psychiatric disorders and related phenotypes has been shown to improve by exploiting genetic correlations among multiple related traits,23 and methods have been extended to incorporate polygenic scores for multiple traits.8

In the current study, we have combined these approaches and implemented them in a newly developed platform, called
GENOSCORES, which contains GWAS data for multiple traits and automates construction of genotypic scores. For polygenic traits with multiple trait-associated loci, such as RA, locus-specific scores can be constructed to examine how genetic heterogeneity of the intermediate trait can influence the trait of interest. Our approach reduces the dimensionality of the prediction task from about 2 million common SNPs to a few hundred or a few thousand genotypic scores, depending on how relevant traits are selected. The score constructions are a type of feature engineering, a task commonly used in machine learning applications.

Understanding the pathogenic mechanisms that initiate and perpetuate RA could give rise to informative biomarkers of prognosis, therapeutic response and toxicity. However, in agreement with earlier studies, we did not find strong predictors of TNFi response among alleles linked to the development of RA. A strength of the current study is the large sample size which allowed us to detect small robust increments in predictive performance. Our methodological approach was to first establish the predictive value for a set of genetic markers using a multivariate model and then to examine univariate associations between each marker and the outcome. Using this approach, we showed that a model including RA scores led to a small robust improvement in prediction, with the regional score at the CD40 locus driving the predictive signal.

Higher RA risk at the CD40 locus, higher CD40 transcription and lower CD40 methylation were associated with better TNFi response. CD40 is a transmembrane protein which belongs to the TNF receptor superfamily, critically important in modulating immune-(auto-immune) responses. CD40 is expressed by B cells and antigen-presenting cells (APCs), whereas CD40L is induced on CD4+ T cells following T-cell antigen receptor (TCR) with major histocompatibility complex (MHC) molecule interaction. Engagement of the CD40–CD40L axis leads to B cell activation, proliferation and (auto)-antibody production, while activation of APCs by CD40L on CD4+ T cells induces upregulation of CD80, CD86, MHC class I and MHC class II, as well as secretion of proinflammatory cytokines such as interleukin (IL)-12, IL-23 and TNF-α.

The risk allele associated with RA is associated with elevated CD40 expression in whole blood. As high CD40 expression is associated with elevated TNF-α production and CD40 and CD40L transcripts are increased in the disease tissue in both early and established disease, it is not surprising that patients with the CD40 risk allele respond better to TNFi therapies.

Overall, if there are genetic loci that predict TNFi response, these are mostly different to the known RA loci. However, we note that patients who receive TNFi therapy are likely to have more severe disease and to have failed on other treatments. It is therefore possible that our study sample has been selected with respect to genetic load for RA, thus limiting the heterogeneity in genetic RA risk profiles compared with a sample of newly diagnosed cases.

The genotypic score for the expression of the ectonucleotidase CD39 on CD4 T cells was inversely associated with TNFi response. The SNPs contributing to this score are in the ENTPD1 gene which encodes CD39. In stratified analyses, the inverse association with response was stronger in the groups receiving TNFi concomitantly with MTX or another non-biologic DMARD compared with the group receiving TNFi monotherapy and was stronger in the group receiving infliximab compared with the groups receiving adalimumab or etanercept. The CIs of the estimated effects among all groups overlapped. This effect on response to TNFi agents is in the opposite direction to the association reported between low expression of CD39 on regulatory T cells and resistance to MTX in RA.

Interpreting the association between drug response and the CD39 trait is difficult both epidemiologically and mechanistically. The reasons for this being that in the UK and most European countries TNFi are usually prescribed only after patients have had a poor response to MTX and, unless of intolerance, always in combination with MTX. Therefore, these patients are likely to represent a selected group; this, together with the almost universal use of combination MTX/TNFi therapy, hinders progress in dissecting the potential mechanisms responsible for the divergence. Nonetheless, as RA is a highly heterogeneous disease, it is plausible to speculate that different cellular and molecular networks may be involved in driving diverse immune/inflammatory responses in different patients to different drugs.

For example, as mentioned above, the poor response to MTX has been associated with low CD39 expression by regulatory T cells, while increased CD39 expression has been reported to be important in the expansion of Th17 cells driven by IL-6 and TGF-β via Stat3 and Gfi-1 transcription factors. In turn, the expansion of Th17 cells has been reported to be associated with incomplete response to TNFi. It remains to be established whether measurements of CD39/ENTPD1 expression or genotype may be useful in the choice of MTX, TNFi agent or concomitant treatment as first-line therapy for patients who need a DMARD.

Using previously reported associations between transcripts and TNFi response to select relevant genes, we have shown evidence that eQTL scores for these genes contain information that predicts TNFi response, even though the proportion of variance explained was low and no single genes associated with response could be identified.

Improved genomic prediction of treatment response requires measuring response more precisely to capture the molecular in addition to the clinical phenotype. The detected associations between genotypic scores and TNFi response were not the same for the two measures of response—change in SJC and change in ESR—suggesting that the two measures reflect different aspects of disease activity affected by TNFi. To derive refined measures of drug response, large data sets with multiple inflammatory biomarkers, joint imaging and clinical variables before and after treatment are needed.

Author affiliations

1Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
2MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
3Arthritis Research UK Centre for Genetics and Genomics, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
4NIHR Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, University of Manchester NHS Foundation Trust, Manchester, UK
5Division of Rheumatology, Department of Medicine, Brigham and Women’s Hospital, Boston, Massachusetts, USA
6Department of Human Genetics, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands
7Department of Orthopedic Surgery, Tokyo Women’s Medical University, Tokyo, Japan
8The Center of Research Excellence in Science and Technology, Japan Science and Technology Agency, Tokyo, Japan
9Department of Rheumatology, School of Medicine, Tokyo Women’s Medical University, Tokyo, Japan
10Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden
11Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
12Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, Alabama, USA
Patient and public involvement statement
This study was conducted as part of the eWaves Stream 2 of the MATURA Consortium (http://www.matura.wi.imul.ac.uk/what_is_matura.php). A patient advisory group was established in 2014 when the MATURA project commenced. The group meets regularly to ensure MATURA strategy is maintaining relevance, accountability and direction by embedding patients and members of the public within the decision making processes. Determine what level of confidence in tests, and what type of tests, would be acceptable to patients and members of the public within the decision making processes. Determine what your research, for instance by producing lay summaries of the results from Matura and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

REFERENCES
1 Vincent FB, Morand EF, Murphy K, et al. Antidrug antibodies (ADA) to tumour necrosis factor (TNF)-specific neutralising agents in chronic inflammatory diseases: a real issue, a clinical perspective. Ann Rheum Dis 2013;72:165–78.
2 Wijbrandts CA, Dijkgraaf MGW, Kraan MC, et al. The clinical response to infliximab in rheumatoid arthritis is in part dependent on pretreatment tumour necrosis factor alpha expression in the synovium. Ann Rheum Dis 2008;67:1139–44.
3 Barton A, Pitzalis C. Stratified medicine in rheumatoid arthritis—the Matura programme. Rheumatology 2017;56:1247–50.
4 Masey J, Plant D, Hyrich K, et al. Genome-wide association study of response to tumour necrosis factor inhibitor therapy in rheumatoid arthritis. Pharmacogenomics J 2018;18:657–64.
5 Cherlin S, Plant D, Taylor JC, et al. Prediction of treatment response in rheumatoid arthritis patients using genome-wide SNP data. Genet Epidemiol 2018;42:754–71.
6 McKeigue P. Sample size requirements for learning to classify with high-dimensional biomarker panels. Stat Methods Med Res 2019;28.
7 Mainiero, Moser G, Chen G-B, et al. Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder. Am J Hum Genet 2015;96:283–94.
8 Kropohl E, Patel H, Newhouse S, et al. Multi-polygenic score approach to trait prediction. Mol Psychiatry 2018;23:1368–74.
9 Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 2014;506:316–81.
10 Roederer M, Quaye L, Mangino M, et al. The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. Cell 2015;161:387–403.
11 Kim TH, Choi SI, Lee YH, et al. Gene expression profile predicting the response to anti-TNF treatment in patients with rheumatoid arthritis; analysis of Geo datasets. Joint Bone Spine 2014;81:325–30.
12 Cui J, Stahl EA, Saeverdottir S, et al. Genome-wide association study and gene expression analysis identifies CD84 as a predictor of response to etanercept therapy in rheumatoid arthritis. PLoS Genet 2013;9:e1003394.
13 Baker JS, Conaghan PG, Smolen JS, et al. Development and validation of modified disease activity scores in rheumatoid arthritis: superior correlation with magnetic resonance imaging-detected synovitis and radiographic progression. Arthritis Rheum 2014;66:799–802.
14 Hensor EMA, McKeigue P, Ling SE, et al. Validity of a two-component imaging-derived disease activity score for improved assessment of synovitis in early rheumatoid arthritis. Rheumatology 2019;49.
15 Westa H-J, Peters MJ, Eisko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 2013;45:1238–43.
16 Gusev A, Ko A, Shu H, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 2016;48:243–52.
17 Gaunt TR, Shibah HA, Hennani M, et al. Systematic identification of genetic influences on methylation across the human life course. Genome Biol 2016;17.
18 Carpenter B, Gelman A, Hoffman MD, et al. Stan : A Probabilistic Programming Language. J Stat Softw 2017;76:1–32.
19 Piironen J, Vehlin A. On the Hyperprior Choice for the Global Shrinkage Parameter in the Horseshoe Prior. In: Singh A, Zhu L, eds. Proceedings of the 20th International Conference on artificial intelligence and statistics. . Fort Lauderdale, FL: USA: PMLR, 2017: 54. 905–13.
20 Carvalho CM, Polson NG, Scott JG. The horseshoe estimator for sparse signals. Biometrika 2010;97:465–80.
21 Dupuis JA, Robert CP. Variable selection in qualitative models via an enotropic proximal power. J. Stat Plan Inference 2003;111:77–94.
22 Stone M. An asymptotic equivalence of choice of model by cross-validation and Akaike’s criterion. Journal of the Royal Statistical Society: Series B 1977;39:44–7.
23 Peres RS, Liew FV, Talbot J, et al. Low expression of CD39 on regulatory T cells as a biomarker for resistance to methotrexate therapy in rheumatoid arthritis. Proc Natl Acad Sci U SA 2015;112:2509–14.
24 Sieberts SK, Zhu T, Garcia-Garcia J, et al. Crowdsourced assessment of common genetic contribution to predicting anti-TNF treatment response in rheumatoid arthritis. Nat Commun 2016;7.
25 Turley P, Walters RK, Maghzian O, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet 2018;50:229–37.

Rheumatoid arthritis

Patient and public involvement statement
This study was conducted as part of the eWaves Stream 2 of the MATURA Consortium (http://www.matura.wi.imul.ac.uk/what_is_matura.php). A patient advisory group was established in 2014 when the MATURA project commenced. The group meets regularly to ensure MATURA strategy is maintaining relevance, accountability and direction by embedding patients and members of the public within the decision making processes. Determine what level of confidence in tests, and what type of tests, would be acceptable to patients and members of the public within the decision making processes. Determine what type of tests, would be acceptable to patients for treatment decisions. Maximise patient recruitment to research studies by increasing awareness through patient groups.Readily obtain patients’ perspective on grant applications related to stratified medicines for RA. Facilitate the dissemination of the results from MATURA research, for instance by producing lay summaries of papers in conjunction with the researchers (http://www.matura.wi.imul.ac.uk/news.php).

Patient consent for publication
Not required.

Ethics approval
This study used anonymised data for human subjects from an international collaboration of 13 studies originally published in PLOS Genetics 2013;9:e1003394. All participants provided informed consent and institutional review board and ethics approvals were in place for each of the studies and described in the publication.
Rheumatoid arthritis

26 McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med 2011;365:2205–19.
27 Jiang X, Askling J, Saevarsdottir S, et al. A genetic risk score composed of rheumatoid arthritis risk alleles, HLA-DRB1 haplotypes, and response to TNFi therapy – results from a Swedish cohort study. Arthritis Res Ther 2016;18.
28 Grewal IS, Flavell RA. The role of CD40 ligand in costimulation and T-cell activation. Immunol Rev 1996;153:85–106.
29 Elgueta R, Benson MJ, de Vries VC, et al. Molecular mechanism and function of CD40/CD40L engagement in the immune system. Immunol Rev 2009;229:152–72.
30 Summers de Luca L, Gommerman JL. Fine-tuning of dendritic cell biology by the TNF superfamily. Nat Rev Immunol 2012;12:339–51.
31 Jacobson EM, Huber AK, Akeno N, et al. A CD40 Kozak sequence polymorphism and susceptibility to antibody-mediated autoimmune conditions: the role of CD40 tissuespecific expression. Genes Immun 2007;8:205–14.
32 Guo Y, Walsh AJ, Fearon U, et al. CD40L-Dependent pathway is active at various stages of rheumatoid arthritis disease progression. J Immunol 2017;198:4490–501.
33 Chalmin F, Mignot G, Bruchard M, et al. STAT3 and Gli-1 transcription factors control Th17 cell immunosuppressive activity via the regulation of ectonucleotidase expression. Immunity 2012;36:362–73.
34 Alzabin S, Abraham SM, Taher TE, et al. Incomplete response of inflammatory arthritis to TNFα blockade is associated with the TH17 pathway. Ann Rheum Dis 2012;71:1741–8.