Myasoedova Elena (Orcid ID: 0000-0003-2006-1436)
Crowson Cynthia S. (Orcid ID: 0000-0001-5847-7475)
Davis III John M (Orcid ID: 0000-0002-9710-8143)
Emery Paul (Orcid ID: 0000-0002-7429-8482)
Matteson Eric L. (Orcid ID: 0000-0002-9866-0124)



# Towards Individualized Prediction of Response to Methotrexate in Early Rheumatoid Arthritis: a Pharmacogenomics-driven Machine Learning Approach

Elena Myasoedova, MD PhD[1,2], ORCID iD: 0000-0003-2006-1436
Arjun P. Athreya, PhD[3]
Cynthia S. Crowson, PhD[1,2], ORCID iD: 0000-0001-5847-7475
John M. Davis III, MD MS[1], ORCID iD: 0000-0002-9710-8143
Kenneth J. Warrington, MD[1]
Robert C. Walchak, BS[4]
Erin Carlson, BS[2]
Krishna R. Kalari, PhD[2]
Tim Bongartz, MD[5]
Paul P. Tak, MD, PhD[6,7], 0000-0002-3532-5409
Ronald F. van Vollenhoven, MD, PhD[6]
Leonid Padyukov, MD, PhD[8] ORCID iD: 0000-0003-2950-5670
Paul Emery, MD[9], ORCID iD: 0000-0002-7429-8482
Ann Morgan, MBChB, PhD[10], ORDID iD: 0000-0003-1109-624X
Liewei Wang, MD, PhD[3]
Richard M. Weinshilboum, MD[3]
Eric L. Matteson, MD MPH[1,2], ORCID 0000-0002-9866-0124,
on behalf of PAMERA consortium

**Key Indexing Terms:** pharmacogenomics, methotrexate, individualized medicine, rheumatoid arthritis

[1] Division of Rheumatology, Department of Internal Medicine, Mayo Clinic, Rochester, Minnesota;

[2] Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota;

[3] Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, Minnesota;

[4] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota

[5] Department of Emergency Medicine, Vanderbilt Medical Center

[6] Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands

[7] Candel Therapeutics, Needham, Massachusetts

[8] Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital, SE-171 76, Stockholm, Sweden

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/acr.24834

[9] Leeds Institute Rheumatic and Musculoskeletal Medicine University Leeds Institute NIHR Biomedical Research Centre, LTHT, Leeds UK; YEAR consortium management team

[10] Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds and NIHR Biomedical Research Centre and Medtech and In vitro Diagnostics Co-operative, Leeds Teaching Hospitals NHS Trust, Leeds, UK

**Funding:** This work was supported by a career development award to Dr. Elena Myasoedova from the Louis V. Gerstner, Jr. Fund at Vanguard Charitable.

The YEAR study was part funded by programme grants from Arthritis Research UK (grant codes 18475 and 18387), the NIHR Leeds Musculoskeletal BRU and Diagnostic Evaluation Co-operative, the British Medical Association (Doris Hillier Award) and the Ann Wilks Charitable Foundation.

**Conflict of interest:** Authors declare no conflict of interest.

**Correspondence to:** Elena Myasoedova, MD, PhD

Mayo Clinic College of Medicine and Science

Division of Rheumatology

200 1st St. SW, Rochester, MN 55905.

Phone: (507) 284-4474

Fax: (507) 284-0564

Email: myasoedova.elena@mayo.edu

**Running head:** Pharmacogenomics in RA

**ABSTRACT**

**Objective.** To test the ability of machine learning (ML) approaches with clinical and genomic biomarkers to predict methotrexate treatment response in patients with early rheumatoid arthritis (RA).

**Methods.** Demographic, clinical and genomic data from 643 patients of European ancestry with early RA (mean age 54 years; 70% female) subdivided into a training (n=336) and validation cohort (n=307) were used. The genomic data comprised 160 single nucleotide polymorphisms (SNPs) previously associated with RA or methotrexate metabolism. Response to methotrexate monotherapy was defined as good or moderate by the European League Against Rheumatism (EULAR) response criteria at 3-month follow-up. Supervised ML methods were trained with 5-repeats and 10-fold cross-validation using the training cohort. Prediction performance was validated in the independent validation cohort.

**Results.** Supervised ML methods combining age, sex, smoking, rheumatoid factor, baseline Disease Activity Score with 28-joint count (DAS28) and 160 SNPs predicted EULAR response at 3 months with the area under the receiver operating curve of 0.84 (p=0.05) in the training cohort and achieved a prediction accuracy of 76% (p=0.05) in the validation cohort (sensitivity 72%, specificity 77%). Intergenic SNPs *rs12446816, rs13385025*, *rs113798271*, and *ATIC (rs2372536)* had variable importance above 60.0 and along with baseline DAS28 were among the top predictors of methotrexate response.

**Conclusion.** Pharmacogenomic biomarkers combined with baseline DAS28 can be useful in predicting response to methotrexate in patients with early RA. Applying ML to

predict treatment response holds promise for guiding effective RA treatment choices, including timely escalation of RA therapies.

**Significance and Innovations**

- Informative predictors of response to treatment with methotrexate in patients with early rheumatoid arthritis (RA) are lacking.
- This study is one of the first to apply machine learning methods integrating clinical and genomic data for individualized prediction of response to methotrexate in patients with early RA.
- Pharmacogenomics biomarkers combined with baseline DAS28 predicted response to methotrexate in patients with early RA more reliably than clinical data alone, with replication in the independent validation cohort.

A treat-to-target approach is a cornerstone of current American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) guidelines for rheumatoid arthritis (RA) treatment, aimed at achieving remission or low disease activity, in order to combat the adverse outcomes in RA (1, 2). Despite the growing armamentarium of RA therapeutics, methotrexate remains the preferred initial disease-modifying antirheumatic drug (DMARD) for RA (1).

While for many patients, methotrexate is the only drug needed to control RA disease activity, 30-40% of patients are non-responders to methotrexate and more than 50% of patients discontinue methotrexate within 3-5 years due to non-response or side effects (3). Composite measures of RA disease activity, i.e. disease activity score with 28 joint count (DAS28) and instruments incorporating change in DAS28 over time in addition to an absolute level of DAS28 (i.e. EULAR response criteria) have been validated for measuring treatment efficacy in RA and reflect a clinically meaningful target of reaching low disease activity or remission (4-6). A 3- to 6-month trial of methotrexate treatment is generally recommended before a decision is made regarding its effectiveness (1). Predicting response to methotrexate early (i.e., at 3-month follow-up) would inform timely escalation of treatment for improved control of RA disease activity, avoiding unnecessary risk of methotrexate-related side-effects in early RA. Clinically useful predictive algorithm(s) effectively identifying patients with early RA who are likely to respond to methotrexate are lacking.

Baseline clinical predictors can reliably identify only a small proportion of patients with RA at high-risk of treatment failure which is insufficient to individualize treatment decisions early in the disease course(7). Importance of genetic factors in the

predisposition to RA and their contribution to high inter-personal heterogeneity of clinical presentation and response to treatments, such as methotrexate, suggests that addition of genetic markers may aid in personalized prediction of methotrexate efficacy. While research in this area is scarce, small studies show that response to methotrexate can be predicted with clinical pharmacogenetic model, although these results have not been externally replicated (8, 9).

We aimed to develop a predictive algorithm for a response to methotrexate based on sociodemographic, clinical, serological and genomic data of patients with early RA using machine learning (ML) methods. We hypothesized that clinical and sociodemographic factors augmented with genomic biomarkers can achieve cross-site replication in prediction of response to methotrexate therapy using the EULAR criteria.

## MATERIALS AND METHODS

### *Patient population*

*Data source.* We leveraged data from our established collaboration between Mayo Clinic and the PhArmacogenetics of Methotrexate in Rheumatoid Arthritis (PAMERA) consortium (10). Data were available on patients with early RA who met 1987 classification criteria for RA (11, 12) and had information on age, sex, race, ever smoking, positivity for rheumatoid factor (RF) and antibody to cyclic citrullinated peptide (anti-CCP), DAS28 with erythrocyte sedimentation rate (ESR) and/ or C-reactive protein (CRP) at baseline and 3 months of treatment, as well as genomic data.

The genome-wide association study (GWAS) single nucleotide polomorphism (SNP) genotyping was performed at RIKEN, Japan, as previously described, using Illumina Human Omni Express + Exome (8). A total of 653711 SNPs passed the quality

control (excluded chromosome Y, mitochondrial DNA, unplaced SNPs, failed SNPs and SNPs with minor allele frequency <0.01). The genomic data comprised numerical genotypes of the 160 SNPs associated with risk of RA and methotrexate metabolism identified during this GWAS with addition of SNPs that were found significant in studies by other groups (Supplemental Table 1). Of them, 35 SNPs were observed, i.e., directly genotyped in our genome-wide PAMERA study (8); 108 SNPs had an imputation quality $R^2$ score of 0.7 or greater. Genotypes were imputed using EZimputer (http://bioinformaticstools.mayo.edu/research/ezimputer/) with the 1000 genomes phase 1 cosmopolitan reference panel (EZimputer uses impute2). For SNPs that were not found in our database, we selected the SNPs that had the highest linkage disequilibrium (LD) and was within 200kb of the original SNP, using the following tool http://grch37.ensembl.org/Homo_sapiens/Tools/LD (Supplemental Table 2). LD R2 for those SNPs were between 0.53 and 1.0. Dprime was between 0.92 and 1.0. Data analysis was done using NCBI build 37.

All patients were treatment-naïve, initiated on methotrexate monotherapy at baseline and completed a 3-month course of methotrexate (PO or SQ). Users of glucocorticoids with a prednisone equivalent dose ≥15 mg/day at 3 months were excluded. Out of the initial sample of 763 patients, patients who were lost to follow-up (n=71), did not have clinical data (n=16), and were not taking methotrexate at 3 months (n=33) were excluded, resulting in 643 patients who were included in the study.

*Clinical outcomes: definition of response.* Responders to methotrexate monotherapy were defined as patients who were on methotrexate at 3-month follow-up and had *good or moderate response by the EULAR response* criteria at 3 months (5).

The following cut-offs were used. Good response: DAS28 is ≤3.2 at 3 months, decrease by >1.2 from baseline; moderate response: DAS28 ≤3.2 and a decrease >0.6 and ≤1.2, or a DAS28 >3.2 but less than or equal to 5.1 plus a decrease >0.6, or a DAS28 >5.1 and a decrease >1.2. Non-responders were defined as patients who were treated with methotrexate for 3 months but no response by the EULAR response criteria at 3 months.

To develop and validate the prediction models, all patients (n=643) were subdivided into a training cohort including patients from Yorkshire, the United Kingdom (n=336) as the largest single-site sample, and a validation cohort of patients recruited elsewhere across Europe (n=307), i.e. Sweden (n=267) and Netherlands (n=40).

*Statistical Methods*

Baseline RA activity data, sociodemographic, clinical, and genomic data were used to predict response to methotrexate treatment. Binary outcomes of response (i.e. *good or moderate response by EULAR criteria*) versus non-response at 3 months were used. Random forests (i.e., *randomForest* R library) were used because of their mathematical ability to handle discrete data (e.g., genotypes) and correlated predictor variables. Random forests also grow multiple decision trees, with the difference that they use samples of data and subsets of the covariates to avoid overfitting. These flexible approaches can capture latent traits and have demonstrated robust predictive capabilities in several clinical applications (13).

The ML strategy is shown in Figure 1. To minimize the effects of overfitting, information leak, and to use all training data, nested cross-validation (nested-CV) with 5 repeats was used to train the classifiers by tuning model hyperparameters, which maximizes the area under the receiver operating curve (AUC) (14). A new random split

of data was created in each repeat, and the nested-CV comprised an outer loop and an inner loop. The outer loop uses fivefold cross-validation to split the data into training data (80% of the data) and validation data (the remaining 20%). The inner loop uses the training data to train the classifier with 10-fold cross-validation, and the trained classifier is tested on the validation data. To minimize the effects of *class imbalance* (i.e., unequal numbers of responders (60%) and non-responders (40%)) in the training data and achieve balanced probabilities of classes for training prediction models, we used the synthetic minority over-sampling (SMOTE) algorithm, which simulates patient profiles of the under-sampled class and up-samples the under-sampled class to ensure that the two classes have equal sizes (15). Following the recommended practice of grid-search to find the optimal number of trees that maximize predictive performance during training only, the tuning parameter (i.e., mTry variable) was set as the square root of the total number of variables used in building the tree, and the number of trees selected was 500, 1000, 1500, and 2000 (16).

Random forests were trained on PAMERA's UK patients' (N = 336) baseline DAS28, RF-positivity, smoking, and sociodemographic factors (Model 1), as well as with the addition of the pharmacogenomics data (Model 2) to predict 3-month response to methotrexate (i.e. *good or moderate response by EULAR criteria*) versus non-response. We then externally validated the trained prediction model using the independent validation cohort of RA patients (N = 307)(17). Subgroup analyses using the same predictors and outcome were performed in a subset of patients who received at least 15 mg/week of methotrexate at 3 months and in a subset of anti-CCP positive patients.

Top predictors (using variable importance) and the highest-achieved AUC are reported for the training models. For the chosen prediction model we report the standard deviation (SD) of AUC derived during cross-validation. For external validation of the trained models, prediction performance is reported using positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, and statistical significance. We also report the importance of predictor variables (using varImp function in R) wherein variables with higher importance scores signify that the mean decrease in prediction accuracy is higher if that particular variable is not used. If the variable importance is zero, it means that the variable is not strongly associated with any of the outcomes. Variable importance from the trained models is tabulated for the top 20 predictors. To study the classifier's classification performance versus the observed event rates in the validation data, the statistical significance of the classifier's accuracy was established using the null information rate (NIR, i.e., the class's prevalence with the largest samples in validation data), which served as a proxy for chance. Calibration was assessed for the validation cohort using calibration plots and the Hosmer-Lemeshow test. Recalibration was performed by re-estimating the intercept and slope for the model prediction in the validation cohort (18).

## RESULTS

### *Clinical characteristics of patients*

Characteristics of the 643 patients are shown in Table 1. Patients from the UK and other Europe were comparable in their age, sex, race and smoking habits. The majority of patients were positive for RF or anti-CCP antibody. EULAR response criteria for good or moderate response at 3 months were met in 55% patients from the UK and 72% of

patients from other Europe. The majority of patients (84% overall) received methotrexate in the dose of 15 mg/week or higher at 3-month follow-up.

*Individualized prediction of response to methotrexate treatment: model performance*

Table 2 shows the models' prediction performance using sociodemographic, clinical and pharmacogenomics data. Age, sex, RF-positivity and baseline DAS28 predicted EULAR response at 3 months with an AUC of 0.54 in the training cohort including all patients, with an accuracy of 55% (p=0.98) in the validation cohort. Supervised ML methods that combined demographics, RF status, baseline DAS28 and 160 SNPs predicted EULAR response at 3 months with an AUC of 0.84 (p=0.05) in the training cohort and achieved prediction accuracies of 76% (p=0.05) in the validation cohort (sensitivity 72%, specificity 77%). The addition of genomic data improved the predictive accuracies of methotrexate response by 21%. However, calibration was poor in the validation set. Following recalibration, the calibration was good for those with low and high probability of response, but calibration was still poor in the mid-range (Figure 2).

In a subset of patients using at least 15 mg/week of methotrexate at 3 months, the performance of the model combining clinical and genomic markers was similar to the original model, with AUC 0.83 and accuracy 74% in the validation cohort (Table 2). In a subset of anti-CCP positive patients, the model combining clinical and genomic markers showed very high sensitivity (98%) and NPV (99%), but overall prediction accuracy (68%) was not statistically significantly better than the NIR (p=0.94).

*Variable importance in clinical and pharmacogenomic models*

In Model 1 (clinical model including all patients), DAS28 and age were the two top predictor variables (importance 100.00 and 89.33, respectively), followed by smoking

status (importance 0.82) and RF-positivity (importance 0.58). Sex was not a significant predictor (importance 0.00).

In Model 2 (clinical and pharmacogenomic model in all patients), the intergenic SNPs *rs12446816*, *rs13385025*, *rs113798271,* and *ATIC* (*rs2372536)* had variable importance above 60.0 and were among the top predictors of methotrexate response. Baseline DAS28 was among the top 10 predictors while age was of lower importance (Table 3). Models including clinical and pharmacogenomic markers in subsets of patients receiving ≥15 mg/week of methotrexate at 3 months and in anti-CCP positive patients had overlap in genomic predictors between each other and with Model 2 (Table 3).

**DISCUSSION**

Despite the growing number of studies reporting promising clinical and genomic markers of response to methotrexate and other anti-rheumatic treatments in RA, the current trial-and-error approach to treatment remains unchallenged due to high clinical heterogeneity of patients caused by genetic polymorphism and uncontrolled variation of environmental risk factors for individuals. Clearly, there is a need to develop a precision medicine approach toward treatment to achieve better results in the populations that are most likely to respond to a given treatment (19). This study is one of the first to apply ML methods to the individualized prediction of response to methotrexate in patients with early RA using sociodemographic, clinical and genomic data, with validation in an independent cohort. Polymorphisms associated with RA susceptibility, disease expression and methotrexate metabolism were among the top-20 predictors of response to methotrexate at 3 months and improved performance of the pharmacogenomic model over the clinical model by 21% in the validation cohort. This supports the value of combining genomic and

clinical data as a promising avenue for development of clinically useful individualized prediction of response to methotrexate therapy.

The statistical significance for prediction performance was borderline. However, consistent with our study hypothesis, the discriminative ability (AUC = 0.84) of the model was better than chance, wherein the NIR was a proxy for chance. Additionally, the model achieved cross-site replication with balanced prediction accuracy of 76% in the validation cohort which reflects the ability to accurately predict both outcomes: response and non-response. While calibration in the validation cohort was poor, it was good except in the middle of the range following recalibration to account for the higher response rate in the validation cohort. The uncertainty regarding model predictions in the middle of the range reflects the heterogeneity of the training and validation cohorts (e.g., higher proportion of patients using at least 15 mg/week of methotrexate in the validation cohort versus training cohort), and indicates partial utility when porting to other cohorts with room for further customization. Despite the improved prediction performance of our model over previous pharmacogenetic-based predictions (8, 9), the direct comparison between ours and this previous study is precluded by the use of different outcome definitions (EULAR response versus DAS44-based response), different timing of outcome assessment (3 versus 6 months) and the use of 160 versus 17 SNPs (8).

While prediction performance of our model in patients using at least 15 mg/week of methotrexate at 3 months was consistent with the model including all patients, the accuracy was lower in the subset of anti-CCP positive patients. This difference can be explained, in part, by specifics of their autoimmune phenotype (i.e. more aggressive disease) and underlying differences in genetic background of anti-CCP positive RA as a

distinct subset of the disease (20, 21). Earlier studies showed that anti-CCP autoantibodies may inform RA treatment response in patients using biologics (e.g. abatacept and rituximab), and our results suggest that differential prediction by anti-CCP antibody status may be possible in methotrexate users, which requires further study (22, 23).

*Genomic predictors*

Among polymorphisms associated with RA susceptibility and disease expression, SNP rs12446816 had the highest predictive importance in our model. This SNP has been significantly ($p<1x10^{-5}$) associated with RA disease activity reflected in tender joint count using 28 joints (TJC28) in our earlier genome-wide association study of response to methotrexate, but was not one of the top SNPs associated with DAS28 response in that study (10). There are currently no known genes annotated for this SNP, but there are several link genes in this locus and the *ERCC4* gene, responsible for DNA repair. The micro-RNA gene, *MIR365A*, is 300K bp from this SNP. The up-regulation of this gene in murine model of collagen induced arthritis was shown to enhance apoptosis and restrain proliferation of synoviocytes in RA (24). Regulatory function of the rs12446816 SNP in human disease remains to be established. A further three of the top significant SNPs have been previously associated with RA disease activity in the same GWAS: rs9910936 in *EFTUD2*; rs113798271, a non-coding SNP (upstream gene *MIR5192* (-28159bp), downstream gene *RN7SL51P* (28686)), and rs77458347 in *COL25A1* (10). Among RA polymorphisms associated with RA susceptibility, rs2234067 in *ETV7* and rs11574914 in *CCL21* emerged among top predictors in our study. Both of these genes have been

previously implicated in pathogenetic pathways related to therapeutic targets and treatment response in RA (25-30).

Similar to the study by Wessels et al. and other reports, polymorphisms implicated in adenosine metabolism encoded by *ATIC* were among the top predictors of response to methotrexate in our study. *ATIC* 347 GG (rs2372536, minor genotype) was previously associated with better response to methotrexate at 3 months compared with 347 GC and CC (31), which is concordant with our results. The minor allele of *ATIC* rs4673993, an intronic non-coding region, has been previously associated with low RA disease activity at 6 months of methotrexate monotherapy in a cross-sectional study (32). It was hypothesized that *ATIC* rs4673993 may play a role in the regulation of gene expression or can be in linkage disequilibrium with other SNPs (e.g. *ATIC* rs2372536) that may influence protein coding or gene expression.

*Clinical and sociodemographic predictors*

The strongest clinical predictor in this study was baseline DAS28. This finding is concordant with previous reports from observational studies and post-hoc analysis of randomized controlled trials (7, 8, 33). Associations of response to methotrexate in RA with age, smoking and RF-positivity in our clinical model are consistent with other reports (7, 9, 34).

Previous studies in early RA from the BeSt (Behandelstrategieen voor Reumatoide Artritis [Treatment strategies for rheumatoid arthritis]) study subcohort and SWEdish PharmacOTherapy (SWEFOT) trial found that female sex was independently predictive of non-response to methotrexate (8, 34). Sex was not found to be an important predictor of response to methotrexate, similar to a study from Nijmegen RA cohort (9).

Concordantly, recent data from well-defined prospective RA cohorts and clinical trials do not support systematic biologic differences between men and women with RA in clinical response, suggesting that definition of metrics of clinical response can bias the associations (35). Indeed, women are less likely than men to achieve remission defined by the DAS28-ESR, but have similar rates of remission or low disease activity defined by DAS28-CRP, CDAI and RAPID3 (35). Use of age- and sex-specific cut-offs for ESR or use of disease activity metrics not including ESR in future studies may help further refine our understanding of predictive value of sex in response to antirheumatic treatments.

Relatively modest predictive importance of sociodemographic and clinical predictors in our study is in line with some earlier reports and reflects on clinical reality wherein clinicians are unable to individualize methotrexate therapy in current practice. For example, in a study from the UK baseline higher naïve T-cell frequency was found to be predictive of early RA remission at 6 months of first therapy with methotrexate, regardless of demographic, clinical or other immunologic parameters, and it was proposed that naïve CD4+ T-cells can be predictive of progression of inflammatory arthritis (36, 37). Taken together with our results, it suggests the role of assayble plasma biomarkers (e.g., genomics) relating to immune system that prove to be vital in individualizing methotrexate therapy in treating early RA.

The results of this study should be considered in the context of the following limitations. While 160 SNPs were included in this study, it is possible that SNPs that were not included may also contribute to the prediction of response to methotrexate. GWAS screening for a larger number of SNPs predictive of methotrexate response will be needed to extend these data. Data on some risk factors that have been previously linked

to lower likelihood of treatment response, including low socioeconomic status, obesity, mental and physical comorbidities would be expected to refine the prediction, but were not available for this study (7, 38, 39).

All patients in this study reported using methotrexate at 3 months. Non-adherence at some point during the follow up in a minority of patients cannot be excluded as shown in previous studies (40). However, we believe it is unlikely to have a major impact on the results. Information on route of methotrexate administration (i.e. oral or subcutaneous), dose escalation, folic acid use, prednisone dosing between baseline and 3 months and creatinine levels was not available. Future studies should evaluate importance of these factors for individualized prediction of response to methotrexate.

Data on anti-CCP antibody were missing for a proportion of subjects and were not included in the main model. However, we present the analysis for a subset of anti-CCP positive patients. This study was not a randomized trial and included patients who have completed 3-month treatment with methotrexate, not accounting for those who were excluded from the study. We had no data on outcomes at 6 months and later time points.

This study included a homogenous population from a racial standpoint (about 98% patients were Caucasians in both cohorts), which is expected to reduce confounding by race but limits the generalizability of the study results to other races and ethnicities.

With less than 20% DAS28-remission rates at 3 months, this study was not adequately powered to predict DAS28-remission. This shortcoming underscores the need for multicenter collaborative studies using larger datasets of patients with early RA on methotrexate monotherapy. Nevertheless, the EULAR response criteria is a validated measure for evaluation of treatment response in RA (4, 5), and good or moderate EULAR

response has been associated with improved functional capacity and less progression in joint damage compared to EULAR non-response (4). EULAR criteria have been used successfully in other recent studies aimed at prediction of response to antirheumatic treatments at 3-month time-point when remission is infrequent (35, 41). Thus, using EULAR criteria at 3 months is well-justified and can be a valuable first step towards individualized prediction of response to methotrexate treatment.

Strengths of the study stem from the high-quality of the uniformly collected data from the PAMERA consortium and advantages of international collaboration, including the unique availability of two geographically independent cohorts of previously treatment-naïve patients with early RA who were treated with methotrexate monotherapy for 3 months with a 3-month follow-up; uniform collection of clinical data at baseline and 3-month follow-up, and availability of sociodemographic and genomic data. While we had no data on outcomes at 6 months, previous studies have shown that in over 75% of patients, the trajectory of response or non-response to methotrexate is consistent between 3 and 6 months (8). Thus, using a 3-month time-point strengthens the study as it may be beneficial for early optimization of treatment.

Our model achieved cross-site (i.e. in the independent cohort) replication in prediction performance, which is an advantage over the previous studies lacking an external validation cohort (8). The model included easily available clinic measures and robust set of genomic markers which improved the performance of the model over the NIR by 4%. The study takes advantage of the ML approach combining sociodemographic, clinical and genomic data. Previous studies by our group have successfully applied a similar approach for prediction of drug response to antidepressants (42). Validation of the

model in a larger sample of patients with early RA, developing individualized prediction of adverse effects with methotrexate treatment and evaluating clinical impact and cost-effectiveness of our predictive model compared to other available models are the next steps in model development, as part of translation of clinical research to clinical practice, according to the Prognosis Research Strategy (PROGRESS) framework (43).

In summary, in patients with early RA pharmacogenomic biomarkers combined with baseline DAS28 predicted response to methotrexate by the EULAR criteria at 3 months more reliably than demographics and DAS28 alone, with replication in the independent validation cohort. These findings suggest that models combining clinical data and pharmacogenomic biomarkers for prediction of methotrexate response hold promise for guiding effective RA treatment choices and timely escalation of RA therapies. Further studies on personalized prediction of response to methotrexate and other anti-rheumatic treatments are warranted to optimize control of RA disease and improve outcomes for patients with RA.

## References


1. Smolen JS, Landewe RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis. 2020;79(6):685-99.
2. 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 Care Res (Hoboken). 2016;68(1):1-25.
3. Aletaha D, Smolen JS**.** Effectiveness profiles and dose dependent retention of traditional disease modifying antirheumatic drugs for rheumatoid arthritis. An observational study. J Rheumatol. 2002;29(8):1631-8.
4. van Gestel AM, Haagsma CJ, van Riel PL**.** Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum. 1998;41(10):1845-50.
5. 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.
6. Fransen J, van Riel PL**.** The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S93-9.
7. Sergeant JC, Hyrich KL, Anderson J, Kopec-Harding K, Hope HF, Symmons DPM, et al. Prediction of primary non-response to methotrexate therapy using demographic, clinical and psychosocial variables: results from the UK Rheumatoid Arthritis Medication Study (RAMS). Arthritis Res Ther. 2018;20(1):147.
8. Wessels JA, van der Kooij SM, le Cessie S, Kievit W, Barerra P, Allaart CF, et al. A clinical pharmacogenetic model to predict the efficacy of methotrexate monotherapy in recent-onset rheumatoid arthritis. Arthritis Rheum. 2007;56(6):1765-75.
9. Fransen J, Kooloos WM, Wessels JA, Huizinga TW, Guchelaar HJ, van Riel PL, et al. Clinical pharmacogenetic model to predict response of MTX monotherapy in patients with established rheumatoid arthritis after DMARD failure. Pharmacogenomics. 2012;13(9):1087-94.
10. Taylor JC, Bongartz T, Massey J, Mifsud B, Spiliopoulou A, Scott IC, et al. Genome-wide association study of response to methotrexate in early rheumatoid arthritis patients. Pharmacogenomics J. 2018;18(4):528-38.
11. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62(9):2569-81.
12. Arnett FC, Edworthy SM, Bloch DA, Mcshane DJ, Fries JF, Cooper NS, et al. The American-Rheumatism-Association 1987 Revised Criteria for the Classification of Rheumatoid-Arthritis. Arthritis and Rheumatism. 1988;31(3):315-24.
13. Fernandez-Delgado M, Cernadas E, Barro S, Amorim D**.** Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? Journal of Machine Learning Research. 2014;15:3133-81.
14. Krumsiek J, Mittelstrass K, Do KT, Stuckler F, Ried J, Adamski J, et al. Gender-specific pathway differences in the human serum metabolome. Metabolomics. 2015;11(6):1815-33.
15. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP**.** SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research. 2002;16(16):321-57.
16. Breiman L**.** Random forests. Machine Learning. 2001;45(1):5-32.
17. Myasoedova E AA, Crowson CS, Weinshilboum R, Wang L, Matteson E**.** Pharmacogenomics-driven individualized prediction of treatment response to methotrexate in patients with rheumatoid arthritis: a machine learning approach. Ann Rheum Dis. 2020;79:594.
18. Pirracchio R, Ranzani OT**.** Recalibrating our prediction models in the ICU: time to move from the abacus to the computer. Intensive Care Med. 2014;40(3):438-41.
19. Wijbrandts CA, Tak PP**.** Prediction of Response to Targeted Treatment in Rheumatoid Arthritis. Mayo Clin Proc. 2017;92(7):1129-43.
20. Miriovsky BJ, Michaud K, Thiele GM, O'Dell JR, Cannon GW, Kerr G, et al. Anti-CCP antibody and rheumatoid factor concentrations predict greater disease activity in men with rheumatoid arthritis. Ann Rheum Dis. 2010;69(7):1292-7.
21. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum. 2006;54(1):38-46.

22. Sokolove J, Schiff M, Fleischmann R, Weinblatt ME, Connolly SE, Johnsen A, et al. Impact of baseline anti-cyclic citrullinated peptide-2 antibody concentration on efficacy outcomes following treatment with subcutaneous abatacept or adalimumab: 2-year results from the AMPLE trial. Ann Rheum Dis. 2016;75(4):709-14.
23. Isaacs JD, Cohen SB, Emery P, Tak PP, Wang J, Lei G, et al. Effect of baseline rheumatoid factor and anticitrullinated peptide antibody serotype on rituximab clinical response: a meta-analysis. Ann Rheum Dis. 2013;72(3):329-36.
24. Wang X, Gong S, Pu D, Hu N, Wang Y, Fan P, et al. Up-regulation of miR-365 promotes the apoptosis and restrains proliferation of synoviocytes through downregulation of IGF1 and the inactivation of the PI3K/AKT/mTOR pathway in mice with rheumatoid arthritis. Int Immunopharmacol. 2020;79:106067.
25. Kane M, Zang TM, Rihn SJ, Zhang F, Kueck T, Alim M, et al. Identification of Interferon-Stimulated Genes with Antiretroviral Activity. Cell Host Microbe. 2016;20(3):392-405.
26. Mavragani CP, La DT, Stohl W, Crow MK**.** Association of the response to tumor necrosis factor antagonists with plasma type I interferon activity and interferon-beta/alpha ratios in rheumatoid arthritis patients: a post hoc analysis of a predominantly Hispanic cohort. Arthritis Rheum. 2010;62(2):392-401.
27. Raterman HG, Vosslamber S, de Ridder S, Nurmohamed MT, Lems WF, Boers M, et al. The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. Arthritis Res Ther. 2012;14(2):R95.
28. Wampler Muskardin T, Vashisht P, Dorschner JM, Jensen MA, Chrabot BS, Kern M, et al. Increased pretreatment serum IFN-beta/alpha ratio predicts non-response to tumour necrosis factor alpha inhibition in rheumatoid arthritis. Ann Rheum Dis. 2016;75(10):1757-62.
29. Muskardin TLW, Niewold TB**.** Type I interferon in rheumatic diseases. Nat Rev Rheumatol. 2018;14(4):214-28.
30. Pickens SR, Chamberlain ND, Volin MV, Pope RM, Talarico NE, Mandelin AM, 2nd, et al. Role of the CCL21 and CCR7 pathways in rheumatoid arthritis angiogenesis. Arthritis Rheum. 2012;64(8):2471-81.
31. Dervieux T, Furst D, Lein DO, Capps R, Smith K, Walsh M, et al. Polyglutamation of methotrexate with common polymorphisms in reduced folate carrier, aminoimidazole carboxamide ribonucleotide transformylase, and thymidylate synthase are associated with methotrexate effects in rheumatoid arthritis. Arthritis Rheum. 2004;50(9):2766-74.
32. 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):613-7.
33. Smolen JS, van Vollenhoven RF, Florentinus S, Chen S, Suboticki JL, Kavanaugh A**.** Predictors of disease activity and structural progression after treatment with adalimumab plus methotrexate or continued methotrexate monotherapy in patients with early rheumatoid arthritis and suboptimal response to methotrexate. Ann Rheum Dis. 2018;77(11):1566-72.
34. Saevarsdottir S, Wallin H, Seddighzadeh M, Ernestam S, Geborek P, Petersson IF, et al. Predictors of response to methotrexate in early DMARD naive rheumatoid arthritis: results from the initial open-label phase of the SWEFOT trial. Ann Rheum Dis. 2011;70(3):469-75.
35. Maynard C, Mikuls TR, Cannon GW, England BR, Conaghan PG, Ostergaard M, et al. Sex Differences in the Achievement of Remission and Low Disease Activity in Rheumatoid Arthritis. Arthritis Care Res (Hoboken). 2020;72(3):326-33.
36. Ponchel F, Goeb V, Parmar R, El-Sherbiny Y, Boissinot M, El Jawhari J, et al. An immunological biomarker to predict MTX response in early RA. Ann Rheum Dis. 2014;73(11):2047-53.
37. Ponchel F, Burska AN, Hunt L, Gul H, Rabin T, Parmar R, et al. T-cell subset abnormalities predict progression along the Inflammatory Arthritis disease continuum: implications for management. Sci Rep. 2020;10(1):3669.
38. Liu Y, Hazlewood GS, Kaplan GG, Eksteen B, Barnabe C**.** Impact of Obesity on Remission and Disease Activity in Rheumatoid Arthritis: A Systematic Review and Meta-Analysis. Arthritis Care Res (Hoboken). 2017;69(2):157-65.
39. Takanashi S, Kaneko Y, Takeuchi T**.** Characteristics of patients with difficult-to-treat rheumatoid arthritis in real-world. Rheumatology (Oxford). 2021.
40. Hope HF, Hyrich KL, Anderson J, Bluett J, Sergeant JC, Barton A, et al. The predictors of and reasons for non-adherence in an observational cohort of patients with rheumatoid arthritis commencing methotrexate. Rheumatology (Oxford). 2020;59(1):213-23.

41. Yoosuf N, Maciejewski M, Ziemek D, Jelinsky SA, Folkersen L, Muller M, et al. Early Prediction of Clinical Response to Anti-TNF Treatment using Multi-omics and Machine Learning in Rheumatoid Arthritis. Rheumatology (Oxford). 2021.
42. Athreya AP, Neavin D, Carrillo-Roa T, Skime M, Biernacka J, Frye MA, et al. Pharmacogenomics-Driven Prediction of Antidepressant Treatment Outcomes: A Machine-Learning Approach With Multi-trial Replication. Clin Pharmacol Ther. 2019.
43. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.

**Figure 1: Machine learning strategy**

**Figure 2. Calibration plot for assessment of the machine learning model in the validation dataset (panel A) and following recalibration (panel B).**

**Table 1. Patient characteristics**

| Variable | UK (n=336) | Other Europe (n=307) | Overall (n=643) |
|---|---|---|---|
| Age, mean (SD) | 58.6 (13.1) | 53.4 (13.8) | 56.1 (13.7) |
| Sex, female, n (%) | 239 (71.1%) | 219 (71.3%) | 458 (71.2%) |
| Race, n (%)<br>- Black | 4 (1.2%) | 4 (1.3%) | 8 (1.2%) |
| - Caucasian | 332 (98.8%) | 299 (97.4%) | 631 (98.1%) |
| - Other | 0 (0.0%) | 4 (1.3%) | 4 (0.6%) |
| Time from baseline to follow-up, days, n (%) | 108.1 (30.3) | 113.2 (28.9) | 110.5 (29.7) |
| Smoking, ever, n (%) | 204 (60.7%) | 181 (59.0%) | 385 (59.9%) |
| Rheumatoid factor positive, n (%)<br>- Missing, n | 233 (71.9%)<br>12 | 198 (65.3%)<br>4 | 431 (68.7%)<br>16 |
| Antibody to cyclic citrullinated peptide, n (%)<br>- Missing, n | 150 (66.4%)<br>110 | 177 (58.6%)<br>5 | 327 (61.9%)<br>115 |
| **Baseline characteristics** | | | |
| DAS28, mean (SD) | 5.6 (1.4) | 5.7 (1.1) | 5.6 (1.3) |
| Methotrexate dose, mg/week; n (%) for each dose | | | |
| Missing, n | 0 | 2 | 2 |
| <15 | 335 (99.7%) | 294 (96.4%) | 629 (98.1%) |
| >/= 15 | 1 (0.3%) | 11 (3.6%) | 12 (1.9%) |
| **Characteristics at 3-month follow-up** | | | |
| DAS28, mean (SD) | 4.4 (1.6) | 4.0 (1.4) | 4.2 (1.5) |
| Methotrexate dose, mg/week; n (%) for each dose | | | |
| <15 | 89 (26.4%) | 15 (4.9%) | 104 (16.1%) |

| | | | |
|---|---|---|---|
| >/= 15 | 247 (73.6%) | 292 (95.1%) | 539 (83.9%) |
| Prednisone use ever during the study, n (%) | 65 (19.3%) | 38 (12.4%) | 103 (16.0%) |
| EULAR response, n (%) | 185 (55.1%) | 221 (72.0%) | 406 (63.1%) |
| Categories of the EULAR response, n (%) | | | |
| good | 68 (20.2%) | 83 (27.0%) | 151 (23.5%) |
| moderate | 117 (34.8%) | 138 (45.0%) | 255 (39.7%) |
| none | 125 (37.2%) | 85 (27.7%) | 210 (32.7%) |
| worsening | 26 (7.7%) | 1 (0.3%) | 27 (4.2%) |
| ACR-EULAR Boolean remission* | 16 (5.0%) | 27 (8.8%) | 43 (6.9%) |
| Remission defined as DAS28<2.6 | 49 (14.6%) | 61 (19.9%) | 110 (17.1%) |

* Remission defined as TJC<=1, SJC<=1, CRP<=10 and PGA<=10

**Table 2. Prediction performance of the models using sociodemographic and clinical data alone, as well as combined with pharmacogenomics data: a) in all patients, b) in patients using at least 15 mg/week of methotrexate and c) in patients positive for antibody to cyclic citrullinated peptide**

| **a) All patients** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Predictors | TRAINING DATA | VALIDATION DATA | | | | | | |
| | AUC (SD) | Accuracy (%) | p-value (NIR) | 95% CI of accuracy | Se (%) | Sp (%) | PPV (%) | NPV (%) |
| Baseline DAS28, rheumatoid factor, sociodemographics (age, sex, smoking) | 0.54 (0.03) | 54.7 | 0.98 (0.72) | 49, 60.4 | 54 | 55 | 31 | 75 |
| Baseline DAS28, rheumatoid factor, sociodemographics (age, sex, smoking) and pharmacogenomics biomarkers | 0.84 (0.02) | 76.2 | 0.05 (0.72) | 71.1,80.9 | 72 | 77 | 55 | 87 |
| **b) Methotrexate Dose Subset** | | | | | | | | |
| Predictors | TRAINING DATA | VALIDATION DATA | | | | | | |
| | AUC (SD) | Accuracy (%) | p-value (NIR) | 95% CI of accuracy | Se (%) | Sp (%) | PPV (%) | NPV (%) |
| Baseline DAS28, rheumatoid factor, sociodemographics (age, sex, smoking) | 0.55 (0.03) | 52.4 | 1 (0.72) | 46.5, 58.3 | 56.7 | 50.7 | 31.3 | 74.5 |
| Baseline DAS28, rheumatoid factor, sociodemographics (age, sex, smoking) and pharmacogenomics biomarkers | 0.83 (00.03) | 74 | 0.06 (0.72) | 68.5, 78.9 | 79.5 | 71.8 | 52.8 | 89.8 |

| c) Anti-CCP Subset | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Predictors | TRAINING DATA | VALIDATION DATA | | | | | | |
| | AUC (SD) | Accuracy (%) | p-value (NIR) | 95% CI of accuracy | Se (%) | Sp (%) | PPV (%) | NPV (%) |
| Baseline DAS28, rheumatoid factor, sociodemographics (age, sex, smoking) | 0.56 (0.05) | 50.85 | 0.99 (0.75) | 43.2, 58.4 | 51.2 | 50.8 | 25 | 76.4 |
| Baseline DAS28, rheumatoid factor, sociodemographics (age, sex, smoking) and pharmacogenomics biomarkers | 0.74 (0.03) | 68.36 | 0.94 (0.75) | 61.0, 75.1 | 97.7 | 59 | 43 | 98.8 |

**Abbreviations:** anti-CCP = antibody to cyclic citrullinated peptide; DAS28 = disease activity score with 28 joints; AUC = area under the receiving operation curve; NIR = null information rate; CI = confidence interval; Se = sensitivity; Sp = specificity; PPV = positive predictive value; NPV = negative predictive value; SD = standard deviation.

**Table 3. Variable importance for prediction of methotrexate response across 3 models**

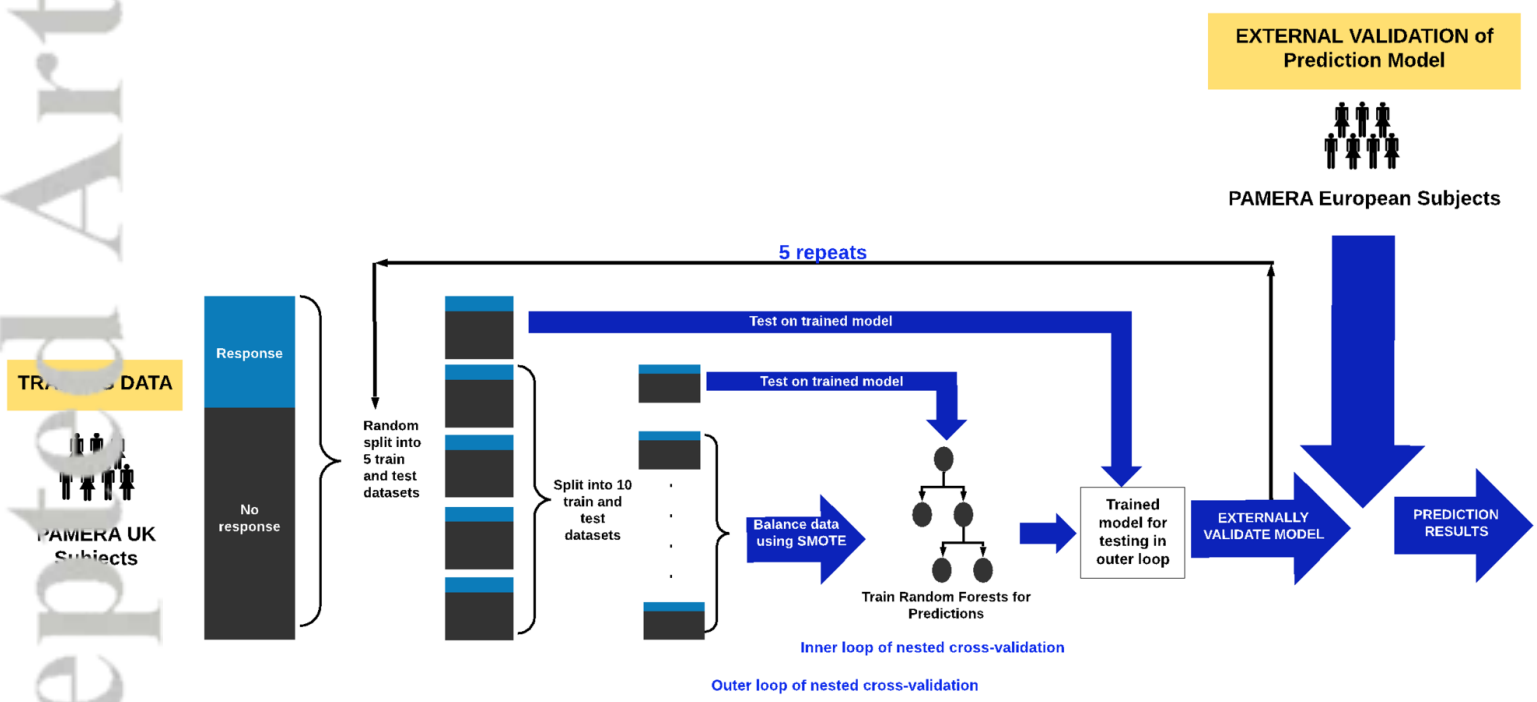
| All patients | | Methotrexate Dose Subset | | Anti-CCP Subset | |
|---|---|---|---|---|---|
| Importance | Variable | Importance | Variable | Importance | Variable |
| 100.00 | rs12446816 | 100.00 | rs2372536 | 100.00 | rs1175813 |
| 70.63 | rs13385025 | 97.99 | rs13385025 | 88.92 | rs71508903 |
| 66.43 | rs113798271 | 90.35 | rs113798271 | 88.02 | rs112502846 |
| 63.86 | rs2372536 | 88.78 | rs10790268 | 87.11 | rs59716545 |
| 57.78 | rs112502846 | 80.71 | rs11258264 | 77.06 | rs11188513 |
| 57.00 | rs10790268 | 75.63 | rs71508903 | 72.92 | DAS28_base |
| 52.61 | DAS28_base | 75.13 | rs12446816 | 72.58 | rs9378815 |
| 50.79 | rs71508903 | 74.42 | rs2639225 | 72.00 | rs1964995 |
| 50.48 | rs2234067 | 74.37 | rs9910936 | 68.79 | rs7071836 |
| 48.44 | rs1964995 | 65.65 | rs1964995 | 64.16 | rs12517451 |
| 48.40 | rs59716545 | 64.32 | rs7563206 | 62.22 | rs10790268 |
| 48.14 | rs34695944 | 63.75 | rs77458347 | 52.79 | rs9603616 |
| 47.76 | rs4673993 | 63.21 | rs9603616 | 48.10 | rs10058818 |
| 46.07 | rs77458347 | 62.56 | rs11889341 | 44.81 | rs314637 |
| 44.99 | rs4272 | 61.98 | rs11933540 | 44.20 | rs10774624 |
| 43.10 | rs11574914 | 60.87 | rs1950897 | 42.89 | rs13385025 |
| 42.47 | rs9603616 | 59.96 | rs3775194 | 42.36 | rs5945173 |
| 42.17 | Age | 59.13 | rs112502846 | 42.28 | rs1950897 |
| 41.00 | rs678347 | 58.71 | rs59716545 | 42.23 | rs73013527 |
| 40.57 | rs9910936 | 57.43 | rs4673993 | 41.70 | rs6903359 |

Common between Anti-CCP and Methotrexate Dose Subsets

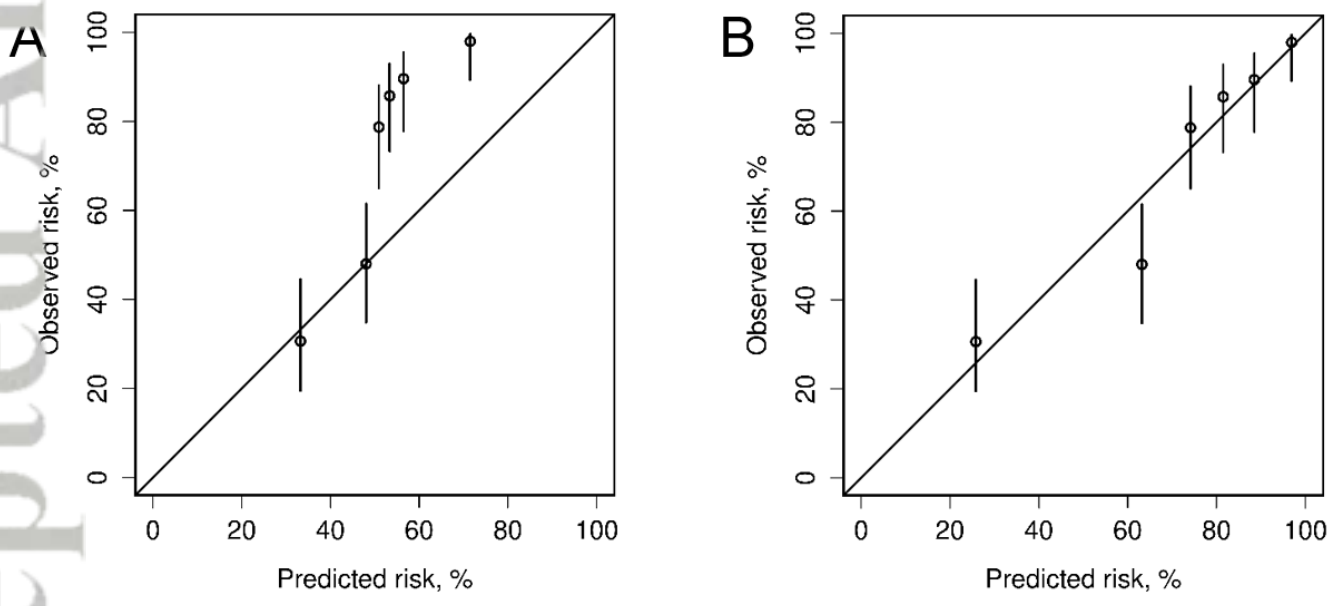
Common between Methotrexate Dose Subsets and All Patients

Common across all 3 groups of patients

**Abbreviations:** DAS28_base = disease activity score with 28 joints at baseline
Anti-CCP = antibody to cyclic citrullinated peptide




ACR_24834_Figure 1.tif




ACR_24834_Figure 2.tif