Journal Club Review of "What Is the Persistence to Methotrexate in Rheumatoid Arthritis, and Does Machine Learning Outperform Hypothesis-Based Approaches to Its Prediction?"

Elizabeth Park

**Objective.** The objectives of this study were to assess the 1-year persistence to methotrexate (MTX) initiated as the first ever conventional synthetic disease-modifying antirheumatic drug in new-onset rheumatoid arthritis (RA) and to investigate the marginal gains and robustness of the results by increasing the number and nature of covariates and by using data-driven, instead of hypothesis-based, methods to predict this persistence.

**Methods.** Through the Swedish Rheumatology Quality Register, linked to other data sources, we identified a cohort of 5475 patients with new-onset RA in 2006-2016 who were starting MTX monotherapy as their first disease-modifying antirheumatic drug. Data on phenotype at diagnosis and demographics were combined with increasingly detailed data on medical disease history and medication use in four increasingly complex data sets (48-4162 covariates). We performed manual model building using logistic regression. We also performed five different machine learning (ML) methods and combined the ML results into an ensemble model. We calculated the area under the receiver operating characteristic curve (AUROC) and made calibration plots. We trained on 90% of the data, and tested the models on a holdout data set.

**Results.** Of the 5475 patients, 3834 (70%) remained on MTX monotherapy 1 year after treatment start. Clinical RA disease activity and baseline characteristics were most strongly associated with the outcome. The best manual model had an AUROC of 0.66 (95% confidence interval [CI] 0.60-0.71). For the ML methods, Lasso regression performed best (AUROC = 0.67; 95% CI 0.62-0.71).

**Conclusion.** Approximately two thirds of patients with early RA who start MTX remain on this therapy 1 year later. Predicting this persistence remains a challenge, whether using hypothesis-based or ML models, and may yet require additional types of data.

https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/acr2.11266

Westerlind H, Maciejewski M, Frisell T, Jelinsky SA, Ziemek D, Askling J. What is the persistence to methotrexate in rheumatoid arthritis, and does machine learning outperform hypothesis-based approaches to its prediction? ACR Open Rheumatol 2021;3:457-463.

**Introduction**

Methotrexate (MTX) remains the first-line disease-modifying antirheumatic drug (DMARD) of choice for patients with newly diagnosed rheumatoid arthritis (RA). It is a cost-effective anchor drug on which other DMARDs are often added. Up to two thirds of patients will initially respond to MTX monotherapy (1), but the rest will require biologic DMARD treatment down the line. Predicting long-term response to MTX is particularly critical to identifying not only those in need of step-up therapy but also refractory patients who are likely to develop erosive, destructive disease. Publications focusing on long-term response are...
somewhat sparse, with one paper (1) citing between 60% and 75% of patients with RA remaining on MTX after 1 to 2 years of initiation. Similarly, two other publications (2,3) cite between 37% and 75% remaining on MTX after 5 years. More recent literature has focused on genetic and clinical predictors of reduced MTX response. A group of studies have (4,5) identified allelic associations, including MTHFR (a gene encoding a key enzyme in folate pathway), solute carrier family 19 member 1 (also known as folate transporter 1), and other proteins involved in purine biosynthesis, with MTX response. Few other studies (1,4,6) have focused on baseline clinical variables (age, sex, disease activity, seropositivity, prior DMARD treatment, etc) as predictors of response. But overall, conflicting results have been reported, with studies unable to be replicated, and most yielding an area under the receiver operating characteristic (AUROC) less than 0.70. Additionally, although individually robust in their methodologies, all of these studies are inherently limited by preselection and manual testing (ie, linear or logistic regression, or scoring systems) of variables (as guided by expert hypotheses), potentially missing other relevant associations. Using machine learning (ML) methods (broadly defined as fitting models algorithmically through adaption of complex patterns in data), may offer an advantage by being able to simultaneously examine thousands of relevant data points to make a prediction.

Methods

Westerlind et al (7) take a new approach to this critical question by 1) combining multiple registers (from the Swedish Rheumatology Quality Register) comprising population/demographic, medication, and insurance data and 2) using diverse ML techniques to produce a global prediction of MTX persistence (defined as remaining on treatment 1 year after initiation without any other DMARDs). This resulted in greater than 5000 patients with new-onset RA starting MTX. By casting a wide net and including all relevant and contributory variables associated with RA diagnosis (dating back up to 10 years prior), the authors amassed more than 4000 unique covariates. Notably, these included American Therapeutic Chemical Classification System and International Classification of Diseases (ICD), 10th Revision codes encompassing various comorbid conditions (acute coronary syndromes, diabetes, heart failure, hepatitis B and C, infections, and other associated rheumatologic conditions, including systemic lupus erythematosus, psoriatic arthritis, and uveitis). These covariate sets were used to predict the main outcome (persistence of MTX at 1 year) by using five different regularized regression/classification ML methods (least absolute shrinkage and selection operator [LASSO], elastic net, support vector machine [SVM], linear kernel, random forest) that were subsequently cross validated five times, (replicating the results on a partitioned data set and allowing for reduction of regression error) and combined as an ensemble to optimize prediction.

The authors then compared the AUROCs resulting from the five ML methods versus manual logistic regression. The accuracy of the ML models was further assessed with calibration plots designed to check for discrepancies between observed and predicted probabilities.

Results

The pooled cohort consisted predominantly of women (68%) with seropositive disease (65% rheumatoid factor–positive) and approximately 10 years of disease duration, similar to other RA clinical cohorts. MTX persistence at 1 year near 70%, but this prevalence dropped to 44% at 3 years (without concurrent use of steroids). Interestingly, equal proportions of patients (13%) reported primary inefficacy of MTX or start of another (additional) DMARD. The best AUROC from ML (LASSO) models was 0.67, equal to that generated from manual logistic regression (0.66). The calibration plot for this LASSO model closely bordered the stand diagonal line of perfect calibration. Moreover, across all four covariate sets, the most important predictors (ie, variables holding the most influence in the ML models) still ended up being baseline clinical variables (such as age, erythrocyte sedimentation rate, C-reactive protein level, Disease Activity Score in 28 joints with C-reactive protein, Health Assessment Questionnaire score, and pain scores). Although when examined by the type of ML model, LASSO and elastic net consistently yielded top predictors derived from ICD codes (eg, time in hospital, diabetes, pregnancy, hepatitis B virus [HBV] diagnosis) rather than clinical variables. One note of caution in interpretation of these findings is that these identified influential variables; although they denote their strong influences in prediction models, they do not necessarily imply explanation of the final outcome (one way or the other). Meaning, although the ICD variable HBV diagnosis played a key role in predicting MTX persistence, it does not necessarily imply that a prior HBV diagnosis led to the avoidance of MTX use and therefore less persistence (although it is reasonable to extrapolate so).

Critique

The main strength of this article is the accurate assessment of MTX persistence (70%) from a merged database enriched with comprehensive and detailed covariates (amassed >4000) associated with RA diagnosis. These covariates spanned wide domains, including conventional clinical characteristics, demographics, medication use, relevant comorbidities, and administrative information. The authors ensured that ML models were robust, accurate, and reliable by applying the trained ML algorithms on separate data held back from training, performing cross-validation, and creating calibration plots (comparing observed vs predicted probabilities). They also maintained
transparency by publishing a clear data pipeline of the chosen ML approaches. The other major/pertinent finding is that baseline clinical factors still emerged as strong predictors of MTX persistence, lending further credence to prior publications and reinforcing the utility of MTX as monotherapy in those with mild disease. Overall, this article adds to the current literature, in which there is a dearth of long-term studies on MTX persistence, particularly those combining clinical and administrative data.

The limitations present are those inherent to using ML. The covariates were not modeled nor fitted prior to testing; this reflects the authors’ intention to first and foremost examine the predictive capacity of this amassed big data rather than prove a preselected hypothesis. Hence the generated AUROC from ML methods ended up being, at best, comparable (<0.70) to those of manual approaches. Additionally, for the main outcome, addition of more complex covariates (such as those drawn from ICD codes) to baseline demographics (age and sex) did not necessarily improve prediction. Although it was not in the scope of this article to offer hypothesis-driven explanations of the covariates (particularly the ICD codes) that predicted MTX persistence, these associations should be investigated in future studies. For instance, time in hospital continued to emerge as an influential variable in various ML models. Considering a separate analysis that explores whether hospitalizations affect cessation or continuation of MTX would be beneficial.

Future studies combining genetic, molecular, and clinical data to predict long-term response and persistence to DMARDs, including biologics, will further advance this field.

**AUTHOR CONTRIBUTIONS**

Dr. Park drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

**REFERENCES**

1. Hider SL, Silman AJ, Thomson W, Lunt M, Bunn D, Symmons DP. Can clinical factors at presentation be used to predict outcome of treatment with methotrexate in patients with early inflammatory polyarthritis? [original article]. Annals of the Rheumatic Diseases 2009;68:57–62.
2. Kinder AJ, Hassell AB, Brand J, Brownfield A, Grove M, Shadforth MF. The treatment of inflammatory arthritis with methotrexate in clinical practice: treatment duration and incidence of adverse drug reactions. Rheumatology (Oxford) 2005;44:61–6.
3. Aletaha D, Smolen JS. Effectiveness profiles and dose dependent retention of traditional disease modifying antirheumatic drugs for rheumatoid arthritis: an observational study. Journal of Rheumatology 2002;29:1631–8.
4. Plant D, Maciejewski M, Smith S, Nair N. Maximising Therapeutic Utility in Rheumatoid Arthritis Consortium, the RAMS Study Group, et al. Profiling of gene expression biomarkers as a classifier of methotrexate nonresponse in patients with rheumatoid arthritis. Arthritis & Rheumatology 2019;71:678–84.
5. Plant D, Wilson AG, Barton A. Genetic and epigenetic predictors of responsiveness to treatment in RA. Nature Reviews Rheumatology 2014;10:329–37.
6. 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 and Rheumatism 2007;56:1765–75.
7. Westerlund H, Maciejewski M, Frisell T, Jelinsky SA, Ziemek D, Asling J. What is the persistence to methotrexate in rheumatoid arthritis, and does machine learning outperform hypothesis-based approaches to its prediction? ACR Open Rheumatol 2021;3:457–63.