Prediction of Non-Response to First-Line Methotrexate Treatment in Rheumatoid Arthritis: A Real-World Data Analysis Using Machine Learning

Zeynep Icten, PhD[1], Kathryn Starzyk, ScM[2], Mark Friedman, MD[1], Joseph Menzin, PhD[1]

[1] Panalgo, Boston MA, USA
[2] OM1, Inc., Boston MA, USA

Background & Objective

- Methotrexate (MTX) is the first choice among disease-modifying drugs (DMARDs) for most rheumatoid arthritis (RA) patients [1].
- However, in a large proportion of patients treated with MTX, disease activity is not adequately controlled [2].
- Timely identification of patients who will not respond well to MTX therapy can protect against progressive and irreversible joint damage. We sought to assess predictors associated with of non-response to MTX using a longitudinal clinical cohort of RA patients and machine learning.

Methods

Data Source

- This study used the OM1 RA Registry (OM1, Inc; Boston, MA) from 01/01/2013 to 03/31/2021

Inclusion and Exclusion Criteria

- The first MTX prescription between 01/01/2014 and 08/01/2020 (Figure 1) was the index date, with no bDMARD/tsDMARD prescribed/administered prior to index
- Patients were 18 years and older at index, had ≥1 Clinical Disease Activity Index (CDAI) from the baseline and post-index, and ≥2 healthcare encounters from the baseline.
- Patients initiating bDMARD/tsDMARDs within 30 days of index date or with other indication(s) for bDMARD/tsDMARDs were excluded.

Study Outcome

- The outcome for prediction was MTX non-response defined as failing to attain remission or low disease activity (CDAI>10) within 4-8 months and/or the initiation of a bDMARD/tsDMARD within 8 months following index date.

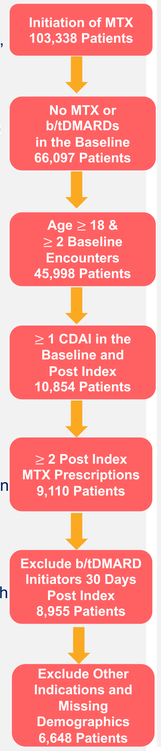
Features

- A total of 250 features from the baseline were identified to be included in subsequent modeling.
- All available ICD-9/10-CM) diagnoses codes were grouped using Clinical Classifications Software (CCS) Single Level schema. Comorbidities with at least a 3% prevalence were kept as features.
- All available NDC and HCPCS medication codes were grouped using the Multum level 3 schema. Medications with at least a 3% prevalence were kept as features.
- All available CPT and HCPCS procedure codes were grouped using the CCS Services and Procedures schema. Procedures with at least a 3% prevalence were kept as features.
- Other features included demographics, indicators of RA-related and all cause healthcare utilization (ER/inpatient visit [any and with RA diagnosis] and other service locations), Charlson Comorbidity Index, latest BMI and CDAI measurements, index MTX prescriber specialty, time between baseline CDAI and index date, year of index date and oral steroids and csDMARDS use.

Predictive Modeling

- Data were partitioned using a 75%/25% split to train/validate and test the predictive model with 3-fold cross validation.
- Models considered included traditional and regularized logistic regression, XGBoost, support vector machine, random forest and feed-forward neural network models. Best model was selected using the area under the ROC curve (AUC) and average precision, Brier score, accuracy, recall, precision, F1 score, negative predictive value and specificity were assessed.
- All analyses were performed using the Instant Health Data (IHD) software (Panalgo LLC, Boston MA, USA).

Figure 1. Study Cohort Attrition



Results

- Among 6,648 eligible RA patients, 3,748 (56.4%) were classified as non-responders (mean age=59.8 years; females=79.1%).
- Baseline demographics stratified by response status are presented in Table 1.
- XGBoost model had the best cross validation AUC (76.4%) improving nearly >4% over the AUC of logistic regression (72.3%).
- The test data AUC of the XGBoost model was 76.03%. Additional performance metrics for the XGBoost model using the test data are presented in Table 2.

Table 1. Baseline Demographics By MTX Response

Study Measures	Response	Non-Response
N	2,900	3,748
Age at Index		
Mean (SD)	63.87 (12.97)	59.81 (13.27)
Median (IQR)	66 (56-73)	61 (51-70)
Female Sex, N (%)	2,096 (72.28%)	2,963 (79.06%)
Region, N (%)		
Midwest	366 (12.62%)	539 (14.38%)
Northeast	520 (17.93%)	595 (15.88%)
South	1,759 (60.66%)	2,304 (61.47%)
West	255 (8.79%)	310 (8.27%)
Race, N (%)		
Black	243 (8.38%)	354 (9.45%)
Caucasian	1,935 (66.72%)	2,370 (63.23%)
Other	73 (2.52%)	93 (2.48%)
Unknown	649 (22.38%)	931 (24.84%)
Payor Type, N (%)		
Commercial	1,516 (52.28%)	2,252 (60.09%)
Medicaid	111 (3.83%)	179 (4.78%)
Medicare	1,273 (43.90%)	1,317 (35.14%)
Ethnicity, N (%)	235 (5.70%)	239 (8.29%)
Hispanic or Latino	127 (4.41%)	208 (5.58%)
Not Hispanic or Latino	1,964 (68.12%)	2,394 (64.27%)
Unknown	792 (27.47%)	1,123 (30.15%)

Figure 2. XGBoost ROC Curve

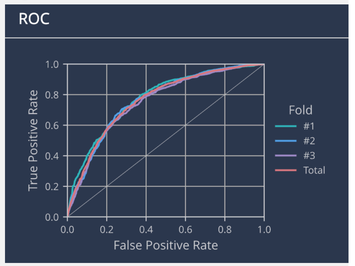


Table 2. XGBoost Test Performance

Metric	Value
AUC	76.03%
Average Precision	78.28%
Brier Score	0.197
Accuracy	70.70%
Recall	71.55%
Precision	74.86%
F1 Score	0.732
Negative Predictive Value	65.94%
Specificity	69.62%

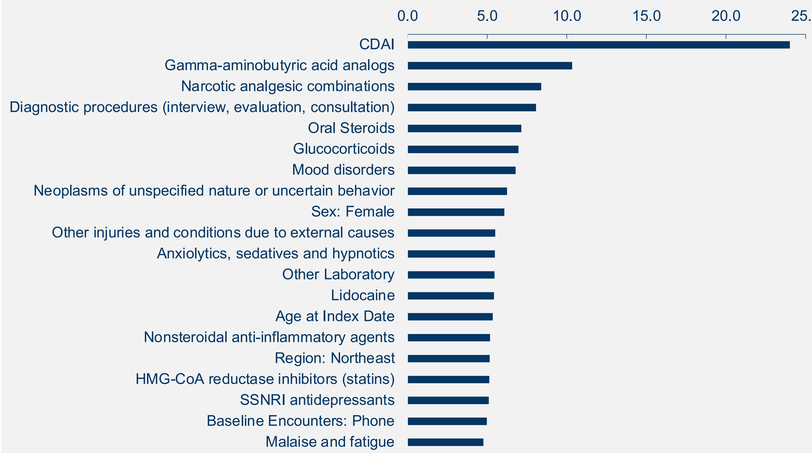
Table 3. XGBoost AUC with and without CDAI Feature

Metric	AUC with CDAI	AUC without CDAI	Decrease in AUC
XGBoost	76.40%	66.80%	9.60%
Random Forest	76.00%	66.89%	9.11%
Regularized Logistic Regression	75.08%	65.69%	9.39%
Support Vector Machine	72.73%	65.47%	7.26%
Logistic Regression	72.32%	63.87%	8.45%
Neural Network	72.16%	65.81%	6.35%

- Regardless of model type, inclusion of CDAI as a feature greatly improved model prediction (Table 3).
- The greatest increase from including CDAI was seen in the XGBoost model.

Results Cont'd.

Figure 3. Top 20 Features of MTX Non-Response According to the Gain Metric



- The XGBoost model included 226 non-zero features. The top 20 features included in the XGBoost model in decreasing order of the gain metric are presented in Figure 3.
- Directionality of the association between the features and outcome was inferred using the SHAP algorithm [3].
- Top predictors of MTX non-response included increasing baseline CDAI, use of GABA analogs, analgesics, oral steroids, glucocorticoids, anxiolytics or sedatives, history of mood disorders, younger age and being female.
- Being from the Northeast appeared protective of MTX non-response.

Conclusion

- Our study identified predictors of MTX non-response with strong predictive accuracy using ML. Further research is needed to better understand the potential role of comorbidities, like mood disorders, and their management on treatment success.
- The inclusion of CDAI, a score not commonly available in claims or general purpose EHR datasets, contributed significantly to the model performance.

References

1. Lopez-Olivo MA, Siddhanamatha HR, Shea B, Tugwell P, Wells GA, Suarez-Almazor ME. Methotrexate for treating rheumatoid arthritis. Cochrane Database Syst Rev. 2014;6:CD000957.
2. Maetzel A, Wong A, Strand V, Tugwell P, Wells G, Bombardier C. Meta-analysis of treatment termination rates among rheumatoid arthritis patients receiving disease-modifying anti-rheumatic drugs. Rheumatology (Oxford). 2000;39(9):975–81. Epub 2000/09/15.
3. Lundberg, Scott M., and Su-In Lee. "A unified approach to interpreting model predictions." Advances in Neural Information Processing Systems. 2017

Presented at ISPOR, May 2022

Funding Source: None