Exploratory Research in Clinical and Social Pharmacy 11 (2023) 100307

Contents lists available at ScienceDirect

Exploratory Research in Clinical and Social Pharmacy

journal homepage: www.elsevier.com/locate/rcsop

Assessing treatment switch among patients with multiple sclerosis: A machine learning approach

Jieni Li [a], Yinan Huang [b], George J. Hutton [c], Rajender R. Aparasu [a,*]

[a] Department of Pharmaceutical Health Outcomes and Policy, College of Pharmacy, University of Houston, TX, USA
[b] Department of Pharmacy Administration, College of Pharmacy, University of Mississippi, Oxford, MS, USA
[c] Baylor College of Medicine, Houston, TX, USA

ARTICLE INFO

Keywords:
Multiple sclerosis
Treatment switching
Real-world evidence
Machine learning

ABSTRACT

Background: Patients with multiple sclerosis (MS) frequently switch their Disease-Modifying Agents (DMA) for effectiveness and safety concerns. This study aimed to develop and compare the random forest (RF) machine learning (ML) model with the logistic regression (LR) model for predicting DMA switching among MS patients.

Methods: This retrospective longitudinal study used the TriNetX data from a federated electronic medical records (EMR) network. Between September 2010 and May 2017, adults (aged ≥18) MS patients with ≥1 DMA prescription were identified, and the earliest DMA date was assigned as the index date. Patients prescribed any DMAs different from their index DMAs were considered as treatment switch. . The RF and LR models were built with 72 baseline characteristics and trained with 70% of the randomly split data after up-sampling. Area Under the Curves (AUC), accuracy, recall, G-measure, and F-1 score were used to evaluate the model performance.

Results: In this study, 7258 MS patients with ≥1 DMA were identified. Within two years, 16% of MS patients switched to a different DMA. The RF model obtained significantly better discrimination than the LR model (AUC = 0.65 vs. 0.63, $p < 0.0001$); however, the RF model had a similar predictive performance to the LR model with respect to F- and G-measures (RF: 72% and 73% vs. LR: 72% and 73%, respectively). The most influential features identified from the RF model were age, type of index medication, and year of index.

Conclusions: Compared to the LR model, RF performed better in predicting DMA switch in MS patients based on AUC measures; however, judged by F- and G-measures, the RF model performed similarly to LR. Further research is needed to understand the role of ML techniques in predicting treatment outcomes for the decision-making process to achieve optimal treatment goals.

1. Introduction

Multiple sclerosis (MS) is a central nervous system inflammatory disease resulting in demyelination and axonal degradation.[1] MS is mainly diagnosed in the reproductive age of women and affects about 0.9 million individuals in the United States (US).[2–4] Diagnosis of MS is associated with functional impairments, lower health-related quality of life (HRQoL), and a significant economic burden.[5–7] Meanwhile, research has found that over 65% of all healthcare-related costs among MS patients were attributed to prescription drugs.[6] There is no specific autoimmune target for MS management; therefore, none of the medications for MS are curative.[8] Accordingly, the

* Corresponding author.
E-mail address: rraparasu@uh.edu (R.R. Aparasu).

https://doi.org/10.1016/j.rcsop.2023.100307
Received 26 April 2023; Received in revised form 8 July 2023; Accepted 9 July 2023
Available online 10 July 2023
2667-2766/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

introduction of disease-modifying agents (DMA) has changed the management landscape by reducing relapse rates, improving patients' HRQoL, and delaying the development of disability.[9–11]

In recent years, treatment options are quickly evolving in MS patients with the introduction of several DMAs.[12] Specifically, the US Food and Drug Administration (FDA) had authorized 21 DMAs to treat MS.[10] With many treatment options for MS, deciding on the initial DMA therapy plan can be challenging.[13] More importantly, each medication failure might result in cumulative neurologic damage, emphasizing the importance of an appropriate therapy plan for newly diagnosed individuals.[14] If the initial DMA is ineffective for MS patients, other DMAs are often prescribed. The findings from the 2019 Milliman MS report suggested that 25.8% of patients switched to another DMA within 12 months after their initial DMA.[15] Several considerations, including tolerability, safety, and effectiveness, influence the decision to change the current DMA for MS patients.[11] Moreover, previous studies based on the traditional regression models identified that age, sex, year of index, relapse, and MS severity were associated with treatment switching among MS patients.[16–18]

Machine learning (ML) algorithms are gaining traction as viable alternatives to traditional regression approaches for prediction and categorization. ML models are increasingly being used in healthcare to generate evidence for improving the quality of care.[19–22] Moreover, ML-derived individual treatment options were proven to generate better outcomes in various diseases.[23–25] However, applications of ML models in MS were limited to predicting the onset, subtypes, and progression of MS with clinical data.[26–29] To the best of our knowledge, ML algorithms have not been applied to examine the treatment-related considerations in MS. Identifying the subpopulation at high risk for treatment switching helps to guide clinical decisions, which is critical for improving health outcomes in MS patients. ML prediction model has the potential to obtain improved accuracy for treatment switching in MS. Hence, this study aimed to develop and validate the ML model to predict DMA switching and further compare the ML model with the traditional regression model for treatment switching among patients with MS.

2. Methods

2.1. Study design and data sources

This retrospective cohort study used the TriNetx, a federated electronic medical record (EMR), from 2009 to 2019 to evaluate treatment switching in MS. TriNetX data consisted of electronic health records for over 84 million patients from 55 healthcare organizations (HCO).[30] In the TriNetX data, both inpatient and outpatient records from the HCOs were included. Overall, the de-identified information in patients' demographic characteristics, diagnoses, procedures, lab tests, vital signs, and prescriptions was combined to assess the utilization of healthcare services. This study was deemed exempt under Category 4 from the Institutional Review Board at the University of Houston.

2.2. Study population

Patients with at least one injectable (including glatiramer acetate and interferon beta) or oral DMA prescription (including fingolimod, dimethyl fumarate, and teriflunomide) were identified during the study period (September 2010–May 2017). Considering the data availability, the DMAs approved before 2017 were considered in this study. Patients' DMA prescription records were identified using RxNorm (produced by the National Library of Medicine) or Healthcare Common Procedure Coding System (HCPCS) codes.[31] The date of the earliest DMA prescription was identified as the index date. Next, patients were selected if they had at least one MS diagnosis within 12 months before or after the index date (identified by the International Classification of Diseases Ninth/Tenth Revision Clinical Modification [ICD-9/10-CM] codes: 340 or G35).[32] Patients with any DMA different from their index drug within 24 months from the index date were considered as treatment switching, and the rest of the patients were flagged as non-switching.

This study did not include patients <18 years old at the index date. Also, patients who were prescribed any DMAs during 12 months washout period were excluded to eliminate the prevalent bias. Moreover, to ensure that patients were continuously involved within the system during the study period, patients with no outpatient visit and no prescription visit within either 12 months pre-index date or 24 months post-index date were excluded. Due to the higher rate of relapse reduction and lower disability progression, infusion DMAs were considered as the high-efficacy treatment. [33,34] Accordingly, patients with any infusion DMA prescriptions (Alemtuzumab, Mitoxantrone, Natalizumab, Ocrelizumab) within 12 months before the index date to the earlier date of the DMA switching or 24 months after the index date were excluded.

2.3. Conceptual framework and measurement of variables

A total of 72 variables were selected for training the ML modelbased on a review of existing literature that identified factors assocated with treatment switching in MS.[13,18,31,35] These factors were further conceptualized based on the Andersen Behavioral Model (ABM) of Health Services Use, which provides a strong conceptual foundation for predicting treatment switching.[36] According to the ABM, healthcare utilization depends on predisposing, enabling, and need factors. Predisposing factors that explain patients' predisposition to use healthcare services include demographic and socioeconomic factors. Enabling factors affect an individual's ability to access healthcare services. Finally, need factors reflect perceived and actual health status.

The following factors were included in this study for model training: (1) Predisposing factors: age, sex, race, and ethnicity; (2) Enabling factors: time-period (year of the index date); and (3) Need factors: Elixhauser Comorbidities, MS-related symptoms (speech symptoms, brainstem symptoms, general symptoms, cerebellar symptoms, difficulty in walking, pyramidal symptoms, sensory

symptoms, bladder/bowel symptoms or sexual dysfunction, visual symptoms, and cerebral/cognitive symptoms), MS symptomatic medication use (fatigue medications, spasticity drugs, antidepressants, analgesics, bladder dysfunction medications, anticonvulsant medications, and cognition medications), relapse drugs use, and healthcare utilization (emergency room visits, inpatient visits, outpatient visits, and use of magnetic resonance imaging [MRI] procedures).[37] To ensure the differences between the approval date of DMAs would not impact the results, the year of the index data was included in the analysis. All of the above factors were measured 12 months before the index date to capture patients' clinically relevant characteristics.

2.4. Machine learning methods

This study used a 70% random sample for the training purpose, in which the model was trained to predict treatment switching. The remaining 30% of the data was used for testing purposes to evaluate the model performance. The non-switching class vastly outnumbered the switching class in the training dataset, with the ratio of switching to non-switching being about 1 to 6, resulting in a class imbalance problem. The imbalanced data might lead to biased prediction toward the non-switching class for machine learning classifiers.[38,39] Thus, the up-sampling method that duplicated random records from the minority class (switching cohort) to generate balanced cohorts was applied to create more balanced data, facilitating the classifiers to deal with the switching and non-switching classes.[40] Another advantage of the up-sampling method is that it can outperform the down-sampling method, which reduces the sample size and increases the risk of overfitting.[41] The up-sampling method was implemented using the R package "caret".[35] All of the analyses were conducted in R version 3.6.0 (R core team, Vienna, Austria).

2.4.1. Random forest (RF) model

The RF model is one of the most popular predictive approaches due to its higher prediction accuracy than other classification and regression models.[42] The tree-based ML method applied the decision tree framework to segregate values of predictors in a series of binary splits. The RF model is a supervised tree-based ensemble learning method that creates many decision trees inferred from bootstrap samples using bagging approaches.[43] Further, using the random feature selection, the ensemble's predictions are aggregated (majority vote) to make the final prediction.[43] Thus, the RF model is more resistant to the over-fitting issue than the classification and regression tree model.[44,45] Compared to traditional regression models, the critical advantage of the RF algorithm that it is a very flexible algorithm that can evaluate more predictor variables that are not limited by model assumptions, such as multicollinearity.[46] This study implemented the RF model with the R package "randomForest".[47]

There were three critical hyperparameters in the RF model that this study focused on: (1) the number of trees (ntree), which indicated the number of trees that the algorithm creates before taking the most votes or averaging the predictions; (2) the number of variables randomly sampled at each split (mtry), which could influence the model's error rates, stability, and accuracy of single trees; (3) the maximal number of leaf nodes for each tree (maxdepth), which impacts the balance and complexity of trees.[48] The optimal values for the above parameters were tuned in the 10-fold cross-validation and then used to decide the final model. Since a large number of trees are built, the RF model will not require a substantial tuning process; in this study, a model consisting of 200 distinct trees (ntree) was generated, sufficiently large enough for out-of-bag error (OOB) to settle down.[49] The default of $\sqrt{n}$ (n means the total number of predictors at each node) is used for the selected features at each node (mtry = 16).[43] Moreover, the maximal number of leaf nodes for each tree was tuned in a grid ranging between 10 and 200, and the optimal number of terminal node trees was obtained at 50. The most influential predictors were identified based on the mean decreased accuracy and mean decreased Gini of the RF model.[50]

2.4.2. Logistic regression (LR) model

The LR model is a traditional regression method for classification, primarily for dichotomous outcomes. The assumptions of the LR model (e.g., linearity assumption, multicollinearity, and homoscedasticity) are challenging to be satisfied because the underlying data-generating model is unknown. In the LR model, the 10-fold cross-validation was applied to get the best value based on the misclassification error to optimize the values for the punishment parameter (λ) while keeping the minimized model deviance. The important predictors for switching were identified based on the statistical significance. In this study, the LR model was implemented with the "glmnet" package.[51]

2.5. Statistical analysis

Using the testing data, this study assessed the LR and the RF model with the area under the receiver operating characteristic curve (AUROC).[52] The Receiver Operating Characteristics (ROC) curve plot shows the sensitivity (true positive rate) as a function of the specificity (false negative) rate for a variety of thresholds. The area under the curve (AUC) for the LR and the RF machine learning classifiers were compared using the 2-sided DeLong test.[53] In addition, considering the imbalanced data, the F-measure and G-measure were introduced for evaluating these models from the testing data.[54] The F-measure has been widely employed in most ML applications, especially for a binary outcome.[55] The F-measure was calculated by the harmonic mean of precision and recall that could integrate precision and recall into a single measurement that accounts for both perspectives.[56] G-measure ($\sqrt{\text{Precision} \times \text{Recall}}$) aimed to evaluate the model performance in predicting the majority and minority groups.[57,58] Also, the G-measure is a measurement that aims to evaluate the model performance by preventing the overfitting of the negative class and underfitting of the positive class in the model.[59] Accordingly, both measures could provide additional justification for unbalanced data. The F- and G- measures for the RF classifier and the LR model were compared based on absolute difference.[57,58] Other performance measurements (accuracy, specificity,

Fig. 1. Study design diagram.

Table 1
Patients baseline characteristics.

	Patients with treatment switching		Patients without treatment switching		*p*-value
	1161 (16.00%)		6097 (84.00%)		
	N	%	N	%	
Predisposing Factors					
Age Category					
Mean, SD	42.83	11.90	47.09	12.12	<0.0001
Younger adults (18–44 years)	646	55.64%	2602	42.68%	<0.0001
Middle-aged adults (45–64 years)	471	40.57%	3001	49.22%	<0.0001
Old adults (≥65 years)	44	3.79%	494	8.10%	<0.0001
Sex					
Female	883	76.06%	4634	76.00%	0.9705
Male	278	23.94%	1463	24.00%	0.9705
Race					
African American	123	10.59%	573	9.40%	0.2045
Whites	955	82.26%	5061	83.01%	0.5333
Others	83	7.15%	463	7.59%	0.5984
Ethnicity					
Hispanic or Latino	50	4.31%	219	3.59%	0.2374
Not Hispanic or Latino	1031	88.80%	5455	89.47%	0.499
Unknown	80	6.89%	423	6.94%	0.9537
Enabling Factor					
Time-period (of index date)					
2010–11	148	12.75%	1108	18.17%	<0.0001
2012–13	391	33.68%	1576	25.85%	<0.0001
2014–15	461	39.71%	2364	38.77%	0.5497
2016–17	161	13.87%	1049	17.21%	0.0052
Index Drug					
Injectable DMA	817	70.37%	3807	62.44%	<0.0001
Interferon beta 1a	253	21.79%	1154	18.93%	0.0237
Interferon beta 1b	55	4.74%	235	3.85%	0.1592
Peginterferon beta	18	1.55%	41	0.67%	0.0023
Glatiramer acetate	491	42.29%	2377	38.99%	0.0348
Oral DMA	344	29.63%	2290	37.56%	<0.0001
Fingolimod	100	8.61%	796	13.06%	<0.0001
Dimethyl fumarate	202	17.40%	1170	19.19%	0.1531
Teriflunomide	42	3.62%	324	5.31%	0.0155
Need Factors					
Elixhauser Comorbidities					
Congestive heart failure	7	0.60%	43	0.71%	0.6992
Cardiac arrhythmia	29	2.50%	218	3.58%	0.0634
Valvular disease	12	1.03%	74	1.21%	0.6032
Pulmonary circulation disorders	3	0.26%	36	0.59%	0.1561
Peripheral vascular disorders	10	0.86%	69	1.13%	0.4158
Hypertension uncomplicated	121	10.42%	781	12.81%	0.0238
Hypertension complicated	8	0.69%	40	0.66%	0.8988
Paralysis	37	3.19%	228	3.74%	0.3575
Other neurological disorders	158	13.61%	705	11.56%	0.0484
Chronic pulmonary disease	59	5.08%	306	5.02%	0.9283
Diabetes uncomplicated	50	4.31%	265	4.35%	0.9514
Diabetes complicated	20	1.72%	84	1.38%	0.3647
Hypothyroidism	68	5.86%	364	5.97%	0.8813
Renal failure	7	0.60%	50	0.82%	0.4423
Peptic ulcer disease, excluding bleeding	7	0.60%	62	1.02%	0.1828
AIDS/HIV	4	0.34%	12	0.20%	0.3253
Lymphoma	0	0.00%	3	0.05%	0.4497
Metastatic cancer	2	0.17%	10	0.16%	0.9494
Liver disease	0	0.00%	20	0.33%	0.0507
Solid tumor without metastasis	14	1.21%	98	1.61%	0.309
Rheumatoid arthritis/collagen related disorders	26	2.24%	127	2.08%	0.7338
Coagulopathy	4	0.34%	62	1.02%	0.027
Obesity	49	4.22%	275	4.51%	0.6611
Weight loss	5	0.43%	71	1.16%	0.0244
Fluid and electrolyte disorders	41	3.53%	234	3.84%	0.6161
Blood loss anemia	0	0.00%	14	0.23%	0.1022
Deficiency anemia	14	1.21%	98	1.61%	0.309
Alcohol abuse	8	0.69%	53	0.87%	0.5376
Drug abuse	7	0.60%	86	1.41%	0.0249
Psychoses	9	0.78%	76	1.25%	0.1713

(*continued on next page*)

Table 1 (*continued*)

	Patients with treatment switching		Patients without treatment switching		*p*-value
	1161 (16.00%)		6097 (84.00%)		
	N	%	N	%	
Depression	233	20.07%	1056	17.32%	0.0247
Elixhauser Comorbidities Score					
Mean, SD	0.22	4.01	0.47	4.67	0.0577
≤0	926	79.76%	4902	80.40%	0.6146
1–4	47	4.05%	277	4.54%	0.4541
≥5	188	16.19%	918	15.06%	0.3234
Other Comorbidities					
Cancer	30	2.58%	202	3.31%	0.1955
Metabolic Disorder					
Thyroid disorders	92	7.92%	439	7.20%	0.3853
Nutritional deficiencies	140	12.06%	840	13.78%	0.1163
Lipid disorders	92	7.92%	568	9.32%	0.1306
Mental Illness					
Anxiety	116	9.99%	582	9.55%	0.6368
Mood disorders	238	20.50%	1084	17.78%	0.0277
Other Neurological disorders					
Paralysis	35	3.01%	227	3.72%	0.2356
Epilepsy, convulsions	37	3.19%	195	3.20%	0.9839
Headache; including migraine	197	16.97%	805	13.20%	0.0007
Eye disorders	133	11.46%	528	8.66%	0.0024
Ear and sense organ disorders	36	3.10%	183	3.00%	0.8561
Circulatory/vascular disorders					
Heart diseases	27	2.33%	220	3.61%	0.0271
Cerebrovascular disease	18	1.55%	117	1.92%	0.3942
Respiratory disorders					
Chronic obstructive pulmonary disease and bronchiectasis	3	0.26%	53	0.87%	0.0292
Genitourinary disorders					
Diseases of the urinary system	166	14.30%	896	14.70%	0.7253
Diseases of skin and subcutaneous tissue	24	2.07%	125	2.05%	0.9701
Musculoskeletal disorders					
Non-traumatic joint disorders	121	10.42%	588	9.64%	0.4132
Spondylosis; intervertebral disc disorders; other back problems	194	16.71%	1076	17.65%	0.4406
Other connective tissue diseases (including fibromyalgia)	320	27.56%	1575	25.83%	0.2187
Ill-defined conditions					
Nausea, vomiting/abdominal	49	4.22%	234	3.84%	0.5371
MS-Related Symptoms					
Bladder/bowel symptoms or sexual dysfunction	182	15.68%	989	16.22%	0.6436
Brainstem symptoms	140	12.06%	603	9.89%	0.0255
Cerebellar symptoms	6	0.52%	59	0.97%	0.135
Cerebral/Cognitive symptoms	133	11.46%	600	9.84%	0.0942
Difficulty walking	130	11.20%	474	7.77%	0.0001
General symptoms	232	19.98%	1252	20.53%	0.6691
Pyramidal symptoms	95	8.18%	522	8.56%	0.6713
Sensory symptoms	235	20.24%	1002	16.43%	0.0016
Speech symptoms	16	1.38%	112	1.84%	0.2763
Visual symptoms	146	12.58%	612	10.04%	0.0096
None of the above	506	43.58%	2883	47.29%	0.0205
MS Severity Score					
Mean, SD	0.93	1.55	0.85	1.51	0.09
MS Symptomatic Medication Use					
Analgesics	382	32.90%	1811	29.70%	0.0296
Antidepressants	342	29.46%	1580	25.91%	0.0122
Bladder dysfunction medications	92	7.92%	554	9.09%	0.2024
Cognition medications	10	0.86%	41	0.67%	0.4801
Fatigue medications	160	13.78%	771	12.65%	0.2889
Other anticonvulsant medications	135	11.63%	570	9.35%	0.0162
Spasticity drugs	388	33.42%	1922	31.52%	0.2037
None of the above	414	35.66%	2472	40.54%	0.0018
Relapse Drugs					
Any corticosteroids	277	23.86%	1103	18.09%	<0.0001
All-Cause Healthcare Utilization					
Any emergency room visits	159	13.70%	762	12.50%	0.2614
# of emergency room visits	0.30	1.10	0.24	0.86	0.0786
Any inpatient visits	101	8.70%	429	7.04%	0.0459
# of inpatient visits	0.16	0.82	0.12	0.63	0.1053
Length of stays	0.61	3.06	0.52	3.51	0.4005

(*continued on next page*)

Table 1 (*continued*)

	Patients with treatment switching		Patients without treatment switching		*p*-value
	1161 (16.00%)		6097 (84.00%)		
	N	%	N	%	
Any outpatient visits	1161	100.00%	6097	100.00%	
# of outpatient visits	10.07	12.01	8.98	12.67	0.0051
MS-related Healthcare Utilization					
Any emergency room visits	95	8.18%	470	7.71%	0.5807
# of emergency room visits	0.14	0.68	0.12	0.51	0.2711
Any inpatient visits	79	6.80%	358	5.87%	0.2207
# of inpatient visits	0.11	0.65	0.08	0.42	0.2103
Length of stays	0.47	2.69	0.41	2.97	0.4662
Any outpatient visits	1040	89.58%	5363	87.96%	0.1173
# of outpatient visits	4.45	5.20	4.01	5.08	0.0077
Any MRI procedures	358	30.84%	1566	25.68%	0.0003
# of MRI procedures	0.44	0.76	0.34	0.68	<0.0001
Vital Sign					
BMI					
Underweight, <18	379	32.64%	1993	32.69%	0.9767
Healthy, 18–25	11	0.95%	62	1.02%	0.828
Overweight, 26–29	267	23.00%	1490	24.44%	0.2935
Obese, ≥30	217	18.69%	1076	17.65%	0.3947
Not recorded	287	24.72%	1476	24.21%	0.7096
Blood Pressure Diastolic					
Normal	539	46.43%	3159	51.81%	0.0008
Abnormal	361	31.09%	1692	27.75%	0.0205
Not recorded	261	22.48%	1246	20.44%	0.1155
Blood Pressure Systolic					
Normal	539	46.43%	3159	51.81%	0.0008
Abnormal	227	19.55%	1043	17.11%	0.0444
Not recorded	395	34.02%	1895	31.08%	0.0481
Lab Test					
Hemoglobin [Mass/volume]					
Normal	576	49.61%	2833	46.47%	0.0489
Abnormal	107	9.22%	599	9.82%	0.5215
Not recorded	478	41.17%	2665	43.71%	0.1096
Leukocytes [#/volume]					
Normal	546	47.03%	2583	42.37%	0.0033
Abnormal	114	9.82%	495	8.12%	0.0555
Not recorded	501	43.15%	3019	49.52%	<0.0001
Hematocrit [Volume Fraction]					
Normal	546	47.03%	2490	40.84%	<0.0001
Abnormal	113	9.73%	549	9.00%	0.4294
Not recorded	502	43.24%	3058	50.16%	<0.0001
Platelets [#/volume]					
Normal	633	54.52%	2882	47.27%	<0.0001
Abnormal	29	2.50%	143	2.35%	0.7543
Not recorded	499	42.98%	3072	50.39%	<0.0001
Creatinine [Mass/volume]					
Normal	511	44.01%	2606	42.74%	0.4225
Abnormal	61	5.25%	372	6.10%	0.2639
Not recorded	589	50.73%	3119	51.16%	0.791
Monocytes [#/volume]					
Normal	592	50.99%	2874	47.14%	0.016
Abnormal	15	1.29%	42	0.69%	0.0329
Not recorded	554	47.72%	3181	52.17%	0.0054

SD: Standard deviation.
DMA: Disease Modifying Agent.
AIDS/ HIV: Acquired Immunodeficiency Syndrome/ Human Immunodeficiency Virus.

precision, and recall) in evaluating the RF and the LR model were also reported based on a confusion matrix using the caret package.[60]

Furthermore, this study compared switching and non-switching cohorts for their baseline characteristics. Additionally, to ensure the power of validation, each baseline characteristic was compared between the training and testing data. All analyses were conducted using server SAS statistical software version 9.4 (SAS Institute Inc., Cary, NC, USA) at a 0.05 level of significance.

3. Results

3.1. Cohort characteristics

This study identified 7258 MS patients who were prescribed ≥1 DMA between September 2010 and May 2017 after applying the inclusion/exclusion criteria. (See Fig. 1) Overall, 1161 (16%) of MS patients switched to a different DMA within two years. From this study sample, 5079 (69.98%) MS patients were selected for the training data, and 2179 (30.02%) MS patients were set into the testing data. Additionally, after applying the up-sampling method, 8566 MS patients (4283 patients in each cohort) were considered in building up the RF and ML model.

3.2. Study population characteristics

Patients who switched to different DMAs were younger than those who did not (mean age: 42.83 vs. 47.09, $p < 0.0001$, Table 1). A higher percentage of patients from the non-switching cohort had oral DMA as their index medication (29.63% vs. 37.56%, $p < 0.0001$). The most prevalent comorbidities among patients who switched were connective tissue diseases (27.56%), mood disorders (20.50%), and depression (20.07%). Patients who switched had a higher proportion of brainstem symptoms (12.06% vs. 9.89%, $p = 0.0255$), walking difficulties (11.20% vs. 7.77%, $p = 0.0001$), sensory symptoms (20.24% vs. 16.43%, $p = 0.0016$), and visual symptoms (12.58% vs. 10.04%, $p = 0.0096$), as well as being more likely to be prescribed at least one MS symptomatic medication (64.34% vs. 59.46%, $p = 0.0018$), and corticosteroids (23.86% vs. 18.09%, $p < 0.0001$). Patients from the switching cohort had more outpatient visits (10.07 vs. 8.98, $p = 0.0051$), and more of them had at least one MRI procedure during the baseline period (30.84% vs. 25.68%, $p = 0.0003$). No significant differences were observed in patients' characteristics between the testing and training datasets (Supplemental Table 1).

3.3. Important predictors of switching

From the RF model, the top five most influential predictors identified in this study were: age, type of index medication, year of index, number of outpatient visits, and Body mass index (BMI) (See Fig. 2). Meanwhile, age, year of index, Elixhauser Comorbidities Score, type of index medication, BMI, comorbidities conditions (cerebral/cognitive symptoms, eye disorders, nutritional deficiencies, hypertension, other neurological disorders, diabetes, thyroid disorders, heart diseases, and spondylosis), lab test (hemoglobin level, platelet count, and monocyte count), and healthcare utilization (inpatient visits, MS-related inpatient, and MS-related outpatient) were statistically significant factors associated with the treatment switching in the LR model.

3.4. Comparison of random forest model and logistic regression

The RF model improved the AUC (0.65 vs. 0.63, $p < 0.0001$) using the testing data compared to the LR model. However, the RF model had a similar G-Measure score as the LR model (0.73 vs. 0.73); the RF model had an almost equal F-Measure score as the LR model (0.72 vs. 0.72). Other performance measurements for both models varied: for RF: the accuracy was 0.61, specificity was 0.63, precision was 0.89, and recall was 60%; for LR, the accuracy was 0.63, specificity was 0.57, precision was 0.87, the recall was 61%. The AUCs observed in training data were comparable to the AUCs observed in testing data in both models. Details of all performance measures can be found in Table 2.

4. Discussion

ML methods have been broadly applied to predict healthcare utilization (e.g., readmissions and inpatient mortality). [23–25,53] In MS, ML models have been applied to predict the onset of MS, MS subtypes, and the progression of MS.[26–29] However, studies utilizing the ML approach to evaluate treatment considerations are very limited. To the best of our knowledge, this study is the first attempt to

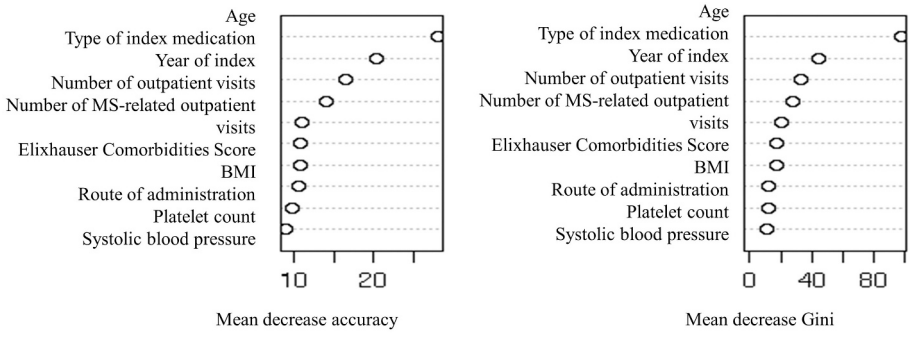

Fig. 2. Top 10 Most Influential Predictors from the Random Forest Model.

Table 2
Model Performance: Random Forest vs. Logistic Regression.

	Test AUC	95% CI	Accuracy	Specificity	Precision	Recall (Sensitivity)	F-1	G-measure	p-value	Train AUC
Logistic Regression	0.6318	0.6007–0.6229	63%	57%	87%	61%	72%	73%	Reference	0.6748
Random Forest	0.6513	0.6204–0.6823	61%	63%	89%	60%	72%	73%	<0.0001	0.7900

AUC: Area under the ROC (Receiver Operating Characteristics) curve.
CI: Confidence interval.

leverage ML methods, namely RF, to predict treatment switching in MS patients. This study developed and compared the RF model with the LR model to predict treatment switching among MS patients using national electronic medical record data. Accurately identifying MS patients at risk of switching to inform targeted intervention may improve MS patients' health outcomes. Literature has supported that ML-derived individual treatment selections were proven to generate better disease outcomes.[23–25] A previous study found that patients who took the RF model suggested epilepsy treatment had considerably higher success rates and lower healthcare utilization than those who took different treatments.[23] Thus, personalized and evidence-based medicine might be achieved using ML approaches for treatment selection and optimizing patient outcomes.

This study adds to existing research regarding the utility and value of ML for treatment considerations. The AUC observed in this study for the RF model was 0.65, an improvement over the AUC of 0.63 observed in the LR model. A previous study found involving the RF model found an AUC of 0.70 for epilepsy treatment, but the treatment change was different from our study, which included any regimen changes or a complete withdrawal.[23] However, the AUC indicated an acceptable model performance distinguishing between switchers and non-switching patients. Additionally, in the case of predicting treatment switching among MS patients, it is preferable to have a greater sensitivity than specificity since identifying patients at risk for treatment switching could assist providers in their treatment decision-making process. As a result, in the interest of healthcare providers who aimed to determine the true positives correctly, a higher sensitivity is preferred , especially when evaluating the model performance.

There is an ongoing debate about the comparative performance of traditional regression approaches compared to ML models.[61] Thus, this study applied several performance metrics to evaluate the relative performance of the RF model vs. the LR model in predicting treatment switching for MS patients. When data is imbalanced, the ROC curve, which plots sensitivity against specificity, was marked as a more robust measurement than accuracy.[62,63] However, misinterpreting specificity might also deceive the AUC when evaluating ML models in imbalanced data.[64] Considering the limitations of AUC, F- and G-measures were identified as more valuable tools because of the ability to detect the true positives correctly.[63] In this study, the RF reported fair scores for F- and G- measures. Although findings showed that the RF offered a notable improvement over LR based on AUC, RF performed similarly to LR judged by F- and G-measures.

Predictive factors are valuable in identifying targets to optimize treatment selection and associated health outcomes in patients with MS. Results from the LR model can help identify the factors and the strength of association with treatment switching among MS patients. Additionally, one of the advantages of applying the RF algorithm is that the model could generate this desired ranking by its relative importance.[65] Among all the factors, the RF model identified the most critical factor in this study was age. This finding is consistent with previous studies that recognize the importance of age on the treatment patterns in MS patients.[13,16–18] A systematic review found that age could impact MS progression, immunosenescence, and DMA selection and engagement.[66] Also, the progression of the disease is observed to be slower among older MS patients than among young and middle-aged MS patients, which might decrease the need for treatment switching among older patients. Furthermore, adverse drug events, such as severe infections, are more commonly observed among the older population.[67] Additionally, the risk-benefit consideration of selecting DMA could be influenced by age-related comorbidities, leading to potential treatment switching among older MS patients.[66]

Moreover, the type of index medication and the year of the index date were observed as the other two top important factors in predicting treatment switching, which could be explained by the introduction of the first oral DMA in 2010.[18] This might also be due to the prescriber's experience and the availability of oral DMAs on the market. Many MS patients switched from injectable to oral DMAs after the launch of oral DMAs.[49,68] Also, evidence from clinical trials and real-world data showed that oral DMAs were associated with reduced relapse rates, delayed MS progression, and higher treatment adherence rates than injectable DMAs.[69–72] In addition to demographic and clinical factors, the RF model revealed the association between BMI and treatment switching among MS patients. Several studies have evaluated the influence of nutrition and diet on the pathophysiology of MS.[73–76] For instance, obese MS patients were reported with higher disease activity than normal or underweight patients.[76,77] Thus, consideration of BMI could add value to making treatment decisions for MS patients based on the literature and findings from this study. Consistent with the results from the RF model, the LR model also found that the abovementioned top important factors were statistically significantly associated with treatment switching among MS patients.

4.1. Strength and limitations

To our knowledge, this is the first study that applied ML methods to examine DMA treatment switching among MS patients. Using a predictive analytics framework to evaluate treatment switching can help manage MS care better and more efficiently. Factors considered in this study involved both demographic and clinical characteristics that are generally richer in the EMR than in claims

data. This study also considered healthcare utilization and lab test when predicting treatment switching. In addition, the study's findings can provide real-world insights into patients' treatment patterns and the factors that could influence treatment switching in MS.

This study, however, also had some limitations that need to be considered. First, medical records from the TriNetX data were mainly gathered from academic hospitals, which may restrict the generalizability of our findings. Secondly, though the rate of treatment switching was low compared to prior research using claims data, the study finding was consistent with previous analyses using electronic medical records.[18] Thirdly, this study developed and compared models in predicting any DMA switching among patients with MS; future studies should compare the treatment switching with different types of DMAs to improve MS care. Fourth, this study evaluated the model performance with an internal validation dataset; more work is needed applying those models with an external dataset to implement the model when predicting treatment switching. Furthermore, several sociodemographic, clinical, and behavioral characteristics, such as insurance status, prescribers, and patients' preferences, were not available in the TriNetX data, which might result in residual confounding and limit the further investigation of our findings. The TriNetX data lacked information on MS activity and duration of MS. Nevertheless; our analysis included the MS severity score, which had previously been validated as a proxy for the disease severity and the disability status.[78,79]

5. Conclusion

The RF model developed in this study provided a better prediction for treatment switching in MS compared to the LR model based on the AUC. The important contributing factors the RF model identified for treatment switching were age, type of index DMA, and time-period of the initial DMA. These factors may help identify targets to optimize treatment selection and associated health outcomes in patients with MS. More research is needed to understand the role and impact of ML algorithms in therapy selection for improving MS care.

Funding

We have no founding support regarding this study.

Disclosures

We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere. A part of the study findings was submitted to the 2022 International Society for Pharmacoeconomics and Outcomes Research (ISPOR) National Conference.

Declaration of Competing Interest

Dr. Aparasu has received research funding from Astellas Inc., Incyte Corp., Gilead, and Novartis Inc. for projects unrelated to the current work. Dr. Hutton reports grants from Biogen, Novartis, MedImmune, Hoffman-LaRoche, E.M.D. Serono, Sanofi, and personal fees from Novartis, Sanofi, Celgene outside the submitted work. The other authors declare no conflicts of interest for this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rcsop.2023.100307.

References

1 Kutzelnigg A, Lassmann H. Pathology of multiple sclerosis and related inflammatory demyelinating diseases. *Handb Clin Neurol.* 2014;122:15–58.
2 Wallin MT, Culpepper WJ, Campbell JD, et al. The prevalence of MS in the United States. *Neurology.* 2019;92(10). https://doi.org/10.1212/WNL.0000000000007035.
3 McGinley MP, Goldschmidt CH, Rae-Grant AD. Diagnosis and treatment of multiple sclerosis. *JAMA.* 2021;325(8). https://doi.org/10.1001/jama.2020.26858.
4 Hunter SF. Overview and diagnosis of multiple sclerosis. *Am J Manag Care.* 2016;22(6 Suppl):s141–s150.
5 Campbell JD, Ghushchyan V, McQueen RB, et al. Burden of multiple sclerosis on direct, indirect costs and quality of life: national US estimates. *Mult Scler Relat Disord.* 2014;3(2):227–236.
6 Earla JR, Thornton JD, Hutton GJ, Aparasu RR. Marginal health care expenditure burden among US civilian noninstitutionalized individuals with multiple sclerosis: 2010-2015. *J Manag Care Spec Pharm.* 2020;26(6):741–749.
7 Li J, Zakeri M, Hutton GJ, Aparasu RR. Health-related quality of life of patients with multiple sclerosis: analysis of ten years of national data. *Mult Scler Relat Disord.* 2022;66, 104019. https://doi.org/10.1016/j.msard.2022.104019.
8 Rae-Grant A, Day GS, Marrie RA, et al. Practice guideline recommendations summary: disease-modifying therapies for adults with multiple sclerosis. *Neurology.* 2018;90(17):777–788. https://doi.org/10.1212/WNL.0000000000005347.
9 Comi G, Radaelli M, Soelberg Sørensen P. Evolving concepts in the treatment of relapsing multiple sclerosis. *Lancet.* 2017;389(10076). https://doi.org/10.1016/S0140-6736(16)32388-1.
10 Elsisi Z, Hincapie AL, Guo JJ. Expenditure, utilization, and cost of specialty drugs for multiple sclerosis in the US Medicaid population, 2008-2018. *Am Health Drug Benefits.* 2020;13(2):74–84.

11 Gajofatto A, Benedetti MD. Treatment strategies for multiple sclerosis: when to start, when to change, when to stop? *World J Clin Cases*. 2015;3(7):545. https://doi.org/10.12998/wjcc.v3.i7.545.
12 Chitnis T, Giovannoni G, Trojano M. Complexity of MS management in the current treatment era. *Neurology*. 2018;90(17). https://doi.org/10.1212/WNL.0000000000005399.
13 Saccà F, Lanzillo R, Signori A, et al. Determinants of therapy switch in multiple sclerosis treatment-naïve patients: a real-life study. *Mult Scler J*. 2019;25(9). https://doi.org/10.1177/1352458518790390.
14 Naismith RT. Multiple sclerosis therapeutic strategies: start safe and effective, reassess early, and escalate if necessary. *Neurol Clin Pract*. 2011;1(1):69–72. https://doi.org/10.1212/CPJ.0b013e31823cc2b0.
15. Milliman Client Report. Multiple Sclerosis: New Perspectives on the Patient Journey - 2019 Update2019. Accessed October 7, 2021 https://us.milliman.com/-/media/milliman/importedfiles/uploadedfiles/insight/2019/ms-patient-journey-2019.ashx; 2019.
16 Freeman L, Kee A, Tian M, Mehta R. Retrospective Claims Analysis of Treatment Patterns, Relapse, Utilization, and Cost Among Patients with Multiple Sclerosis Initiating Second-Line Disease-Modifying Therapy. *Drugs Real World Outcomes*. 2021 Dec;8(4):497–508. https://doi.org/10.1007/s40801-021-00251-w.
17 Desai RJ, Mahesri M, Gagne JJ, et al. Utilization patterns of Oral disease-modifying drugs in commercially insured patients with multiple sclerosis. *J Manag Care Spec Pharm*. 2019;25(1). https://doi.org/10.18553/jmcp.2019.25.1.113.
18 Li J, Chikermane SG, Earla JR, Hutton GJ, Aparasu RR. Factors associated with switching from injectable to oral disease modifying agents among patients with multiple sclerosis. *Mult Scler Relat Disord*. 2022;60, 103703. https://doi.org/10.1016/j.msard.2022.103703.
19 Ling M, Tao X, Ma S, et al. Predictive value of intraoperative facial motor evoked potentials in vestibular schwannoma surgery under 2 anesthesia protocols. *World Neurosurg*. 2018;111:e267–e276.
20 Liang Z, Zhang G, Huang JX, Hu QV. Deep learning for healthcare decision making with EMRs. In: *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE; 2014:556–559.
21 Manogaran G, Lopez D. A survey of big data architectures and machine learning algorithms in healthcare. *Int J Biomed Eng Technol*. 2017;25(2–4):182–211.
22 Buchlak QD, Esmaili N, Leveque JC, et al. Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. *Neurosurg Rev*. 2020;43(5):1235–1253.
23 Devinsky O, Dilley C, Ozery-Flato M, et al. Changing the approach to treatment choice in epilepsy using big data. *Epilepsy Behav*. 2016;56:32–37. https://doi.org/10.1016/j.yebeh.2015.12.039.
24 Wu CS, Luedtke AR, Sadikova E, et al. Development and validation of a machine learning individualized treatment rule in first-episode schizophrenia. *JAMA Netw Open*. 2020;3(2), e1921660. https://doi.org/10.1001/jamanetworkopen.2019.21660.
25 Senders JT, Staples PC, Karhade AV, et al. Machine learning and neurosurgical outcome prediction: a systematic review. *World Neurosurg*. 2018;109:476–486.e1. https://doi.org/10.1016/j.wneu.2017.09.149.
26 Seccia R, Romano S, Salvetti M, Crisanti A, Palagi L, Grassi F. Machine learning use for prognostic purposes in multiple sclerosis. *Life*. 2021;11(2):122. https://doi.org/10.3390/life11020122.
27 Darvishi S, Hamidi O, Poorolajal J. Prediction of multiple sclerosis disease using machine learning classifiers: a comparative study. *J Prev Med Hyg*. 2021;62(1):E192–E199. https://doi.org/10.15167/2421-4248/jpmh2021.62.1.1651.
28 Eshaghi A, Young AL, Wijeratne PA, et al. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nat Commun*. 2021;12(1):2078. https://doi.org/10.1038/s41467-021-22265-2.
29 Vázquez-Marrufo M, Sarrias-Arrabal E, García-Torres M, Martín-Clemente R, Izquierdo G. A systematic review of the application of machine-learning algorithms in multiple sclerosis. *Neurologia (Engl Ed)*. 2021 Feb 3. https://doi.org/10.1016/j.nrl.2020.10.017. S0213-4853(20)30431-X. English, Spanish.
30 Stapff M, Hilderbrand S. First-line treatment of essential hypertension: a real-world analysis across four antihypertensive treatment classes. *J Clin Hypertens*. 2019;21(5). https://doi.org/10.1111/jch.13531.
31 Ontaneda D, Nicholas J, Carraro M, et al. Comparative effectiveness of dimethyl fumarate versus fingolimod and teriflunomide among MS patients switching from first-generation platform therapies in the US. *Mult Scler Relat Disord*. 2019:27. https://doi.org/10.1016/j.msard.2018.09.038.
32 Culpepper WJ, Marrie RA, Langer-Gould A, et al. Validation of an algorithm for identifying MS cases in administrative health claims datasets. *Neurology*. 2019;92(10):e1016–e1028. https://doi.org/10.1212/WNL.0000000000007043.
33 Giovannoni G, Butzkueven H, Dhib-Jalbut S, et al. Brain health: time matters in multiple sclerosis. *Mult Scler Relat Disord*. 2016;9:S5–S48. https://doi.org/10.1016/j.msard.2016.07.003.
34 Vollmer TL, Nair K, v., Williams IM, Alvarez E.. Multiple sclerosis phenotypes as a continuum. *Neurol Clin Pract*. 2021;11(4):342–351. https://doi.org/10.1212/CPJ.0000000000001045.
35 Kuhn M. The caret package. In: *R Foundation for Statistical Computing, Vienna, Austria*; 2023. URL https://cran.r-project.org/web/packages/caret/caret.pdf. Published March 21, 2023.
36 Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav*. 1995 Mar;36(1):1–10.
37 Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1). https://doi.org/10.1097/00005650-199801000-00004.
38 Blagus R, Lusa L. Class prediction for high-dimensional class-imbalanced data. *BMC Bioinformatics*. 2010;11(1):1–17.
39 Kamalov F, Thabtah F, Leung HH. Feature Selection in Imbalanced Data. *Ann. Data. Sci.*; 2022. https://doi.org/10.1007/s40745-021-00366-5.
40 Provost F. Machine learning from imbalanced data sets 101. In: *Proceedings of the AAAI'2000 Workshop on Imbalanced Data Sets*vol. 68.
41 van Smeden M, Moons KG, de Groot JA, et al. Sample size for binary logistic prediction models: beyond events per variable criteria. *Stat Methods Med Res*. 2019;28(8):2455–2474. https://doi.org/10.1177/0962280218784726.
42 Verikas A, Gelzinis A, Bacauskiene M. Mining data with random forests: a survey and results of new tests. *Pattern Recogn*. 2011;44(2):330–349.
43 Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324.
44 Acion L, Kelmansky D, van der Laan M, Sahker E, Jones D, Arndt S. Use of a machine learning framework to predict substance use disorder treatment success. *PloS One*. 2017;12(4), e0175383. https://doi.org/10.1371/journal.pone.0175383.
45 Chen C, Liaw A, Breiman L. *Using Random Forest to Learn Imbalanced Data*110(1−12). Berkeley: University of California; 2004:24.
46 Deo RC, Nallamothu BK. Learning about machine learning: the promise and pitfalls of big data and the electronic health record. *Circ Cardiovasc Qual Outcomes*. 2016;9(6):618–620.
47 Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2(3):18–22.
48 Probst P, Wright MN, Boulesteix A. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2019;9(3), e1301.
49 Desai RJ, Wang S, v., Vaduganathan M, Evers T, Schneeweiss S.. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. *JAMA Netw Open*. 2020;3(1), e1918962. https://doi.org/10.1001/jamanetworkopen.2019.18962.
50 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837. https://doi.org/10.2307/2531595.
51 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1.
52 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837. https://doi.org/10.2307/2531595.
53 Delong ER, Delong DM, Clarke-Pearson DL. *Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach*. 44. 1988. https://about.jstor.org/terms.
54 Dugan JB, Bavuso SJ, Boyd MA. Fault trees and sequence dependencies. In: *Annual Proceedings on Reliability and Maintainability Symposium*. IEEE; 1990:286–293.
55 Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*. 2020;21(1):6. https://doi.org/10.1186/s12864-019-6413-7.

56 Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;15(1):1–28.
57 Kubat M, Holte RC, Matwin S. Machine learning for the detection of oil spills in satellite radar images. *Mach Learn*. 1998;30(2):195–215.
58 Sokolova M, Japkowicz N, Szpakowicz S. *Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation*. 2006:1015–1021. https://doi.org/10.1007/11941439_114.
59 Josephine SA. *Predictive accuracy: a misleading performance measure for highly imbalanced data classified negative*. SAS Global Forum; 2017.
60 Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. 2008;28:1–26.
61 Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12–22.
62 Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. In: *Proceedings of the 23rd International Conference on Machine Learning*. 2006: 233–240.
63 Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS One*. 2015;10(3), e0118432.
64 He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng*. 2009;21(9):1263–1284.
65 Wadekar AS. Understanding opioid use disorder (OUD) using tree-based classifiers. *Drug Alcohol Depend*. 2020;208, 107839. https://doi.org/10.1016/j.drugalcdep.2020.107839.
66 Jakimovski D, Eckert SP, Zivadinov R, Weinstock-Guttman B. Considering patient age when treating multiple sclerosis across the adult lifespan. *Expert Rev Neurother*. 2021;21(3):353–364. https://doi.org/10.1080/14737175.2021.1886082.
67 Patti F, Penaherrera JN, Zieger L, Wicklein EM. Clinical characteristics of middle-aged and older patients with MS treated with interferon beta-1b: post-hoc analysis of a 2-year, prospective, international, observational study. *BMC Neurol*. 2021;21(1):324. https://doi.org/10.1186/s12883-021-02347-w.
68 Earla JR, Paranjpe R, Kachru N, Hutton GJ, Aparasu RR. Use of disease modifying agents in patients with multiple sclerosis: analysis of ten years of national data. *Res Social Adm Pharm*. 2020;16(12). https://doi.org/10.1016/j.sapharm.2020.02.016.
69 Earla JR, Hutton GJ, Thornton JD, Chen H, Johnson ML, Aparasu RR. Comparative adherence trajectories of Oral Fingolimod and injectable disease modifying agents in multiple sclerosis. *Patient Prefer Adherence*. 2020;14. https://doi.org/10.2147/PPA.S270557.
70 Boster A, Nicholas J, Wu N, et al. Comparative effectiveness research of disease-modifying therapies for the Management of Multiple Sclerosis: analysis of a large health insurance claims database. *Neurol Ther*. 2017;6(1). https://doi.org/10.1007/s40120-017-0064-x.
71 English C, Aloi JJ. New FDA-approved disease-modifying therapies for multiple sclerosis. *Clin Ther*. 2015;37(4). https://doi.org/10.1016/j.clinthera.2015.03.001.
72 Scolding N, Barnes D, Cader S, et al. Association of British Neurologists: revised (2015) guidelines for prescribing disease-modifying treatments in multiple sclerosis. *Pract Neurol*. 2015;15(4). https://doi.org/10.1136/practneurol-2015-001139.
73 Markianos M, Evangelopoulos ME, Koutsis G, Davaki P, Sfagos C. Body mass index in multiple sclerosis: associations with CSF neurotransmitter metabolite levels. *Int Sch Res Notices*. 2013:2013.
74 Gianfrancesco MA, Barcellos LF. Obesity and multiple sclerosis susceptibility: a review. *J Neurol Neuromed*. 2016;1(7):1.
75 Mowry EM, Azevedo CJ, McCulloch CE, et al. Body mass index, but not vitamin D status, is associated with brain volume change in MS. *Neurology*. 2018;91(24): e2256–e2264.
76 Kvistad SS, Myhr KM, Holmøy T, et al. Body mass index influence interferon-beta treatment response in multiple sclerosis. *J Neuroimmunol*. 2015;288:92–97.
77 Dardiotis E, Tsouris Z, Aslanidou P, et al. Body mass index in patients with multiple sclerosis: a meta-analysis. *Neurol Res*. 2019;41(9):836–846. https://doi.org/10.1080/01616412.2019.1622873.
78 Nicholas J, Ontaneda D, Carraro M, et al. Development of an algorithm to identify multiple sclerosis (MS) disease severity based on healthcare costs in a US Administrative Claims Database (P2.052). *Neurology*. 2017;88(16 Supplement). https://n.neurology.org/content/88/16_Supplement/P2.052.
79 Toliver JC, Barner JC, Lawson KA, Rascati KL. Use of a claims-based algorithm to estimate disease severity in the multiple sclerosis Medicare population. *Mult Scler Relat Disord*. 2021;49, 102741. https://doi.org/10.1016/j.msard.2021.102741.