Prediction of Rare Disease Medication Transition Patterns Based on An Interpretable Machine Learning Model

RWD106

Jingyi Li[1], Danyang Wei[1], Xuanqi Qiao[1] , Hongfei Gu[2] , Min Hu[1]

[1]Department of Health Economics, School of Public Health, Fudan University, Shanghai, China

[2]Hongmian Cancers and Rare Disorders charity Foundation, Guangzhou, China

BACKGROUNDS

- Patients with rare diseases such as Multiple Sclerosis (MS) have long faced challenges in medication access. Their medication behaviors are shaped by a combination of factors including changes in health insurance coverage policies, income levels, and health status.
- Existing studies, mostly based on cross-sectional data, are unable to capture the dynamic influence mechanisms of the factors.
- Hence, tracking and predicting how these behaviors evolve over time is essential for designing targeted policies that effectively address access barriers.

OBJECTIVE

- To predict patterns of medication transitions and elucidate their underlying drivers among MS patients by developing an interpretable machine learning model with dynamic, longitudinal data, with the ultimate goal of informing policy recommendations to enhance medication accessibility and optimize treatment regimens for rare disease patients.

METHODS

- Data Source: A two-wave prospective longitudinal survey (2022 baseline & 2025 follow-up) of MS patients in China, tracking dynamic changes in insurance, income, health status, and medication channels.
- Data Processing: Longitudinal data matching and cleaning were performed in R. Missing data were handled using Multiple Imputation.
- Statistical & Machine Learning Analysis
  - Descriptive statistics (e.g., means, standard deviations) summarized baseline characteristics.
  - The primary outcome was a four-category **medication transition pattern**: (1) persistent treatment, (2) treatment initiation, (3) treatment discontinuation, and (4) never treated.
  - Predictive models (LightGBM, Random Forest, Logistic Regression) were developed to predict these medication transition patterns. The optimal model was selected based on validation set performance.
  - SHAP analysis was applied to interpret the best model and identify key drivers.

RESULTS

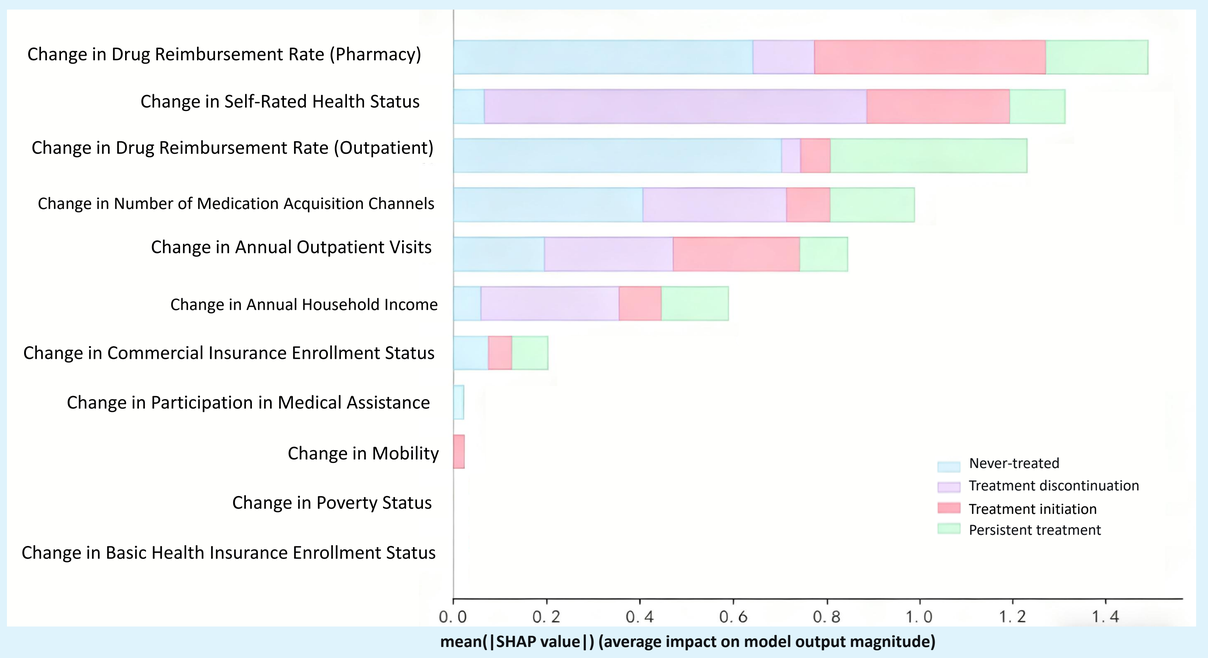
- A total of 455 MS patients were included in this study. The majority of the participants were female (67%). Other baseline characteristics of the study population are summarized in **Table 1**.
- In predicting medication transitions, the LightGBM algorithm achieved superior performance (AUC = 0.84), significantly outperforming both Logistic Regression (AUC = 0.67) and Random Forest (AUC = 0.70) models.
- Drivers of Medication Transitions via SHAP Analysis
  - Treatment Initiation: Increased drug reimbursement was the dominant predictor for initiating therapy (|SHAP| = 0.4988), confirming that lower out-of-pocket costs drive drug utilization.
  - Treatment Discontinuation: Improved self-rated health was the primary driver for discontinuation (|SHAP| = 0.8151), linking symptom relief to reduced adherence.
  - Inadequate Supplemental Coverage: Changes in medical assistance had a minimal predictive effect (|SHAP| < 0.1), calling for enhanced specialized coverage beyond basic insurance.

Table 1 Baseline Characteristics and medication transition patterns of the Study Cohort

Characteristic	2022 (n=455)	2025 (n=455)
Age (years), Mean ± SD	34.5±9.9	36.8±10.2
EDSS score, Mean ± SD	3.41±2.15	3.28±2.09
Health utility score, Mean ± SD	0.73 ± 0.27	0.79 ± 0.25
Sex, n (%)		
Male	150 (33.0)	-
Female	305 (67.0)	-
MS type, n (%)		
RRMS	305 (67.0)	352 (77.4)
Not active SPMS	4 (0.9)	3 (0.7)
Active SPMS	10 (2.2)	7 (1.5)
SPMS (type unknown)	20 (4.4)	18 (4.0)
PPMS	17 (3.7)	10 (2.2)
CIS	9 (2.0)	7 (1.5)
Unknown	90 (19.8)	58 (12.7)
Basic health insurance, n (%)		
Urban Employee Basic Medical Insurance	291 (64.0)	326 (71.6)
Urban and Rural Resident Basic Medical Insurance	144 (31.6)	117 (25.7)
Uninsured	20 (4.4)	12 (2.6)
Commercial health insurance, n (%)		
Enrolled	53 (11.6)	84 (18.5)
Not enrolled	402 (88.4)	371 (81.5)
Annual household income, n (%)		
< ¥50,000	283 (62.2)	160 (35.2)
¥50,000 - ¥200,000	152 (33.4)	265 (58.2)
> ¥200,000	20 (4.4)	30 (6.6)
Medication Transition Patterns ,n(%)		
persistent treatment	267(58.7)	
treatment initiation	69 (15.2)	
treatment discontinuation	25 (5.5)	
never treated	84 (18.5)	

Notes: RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis ; PPMS, primary progressive multiple sclerosis; CIS, clinically isolated syndrome

Figure 1 Summary of SHAP Values for Predictive Features



CONCLUSIONS

- This study demonstrates the potential of machine learning approaches in predicting medication behavior changes among patients with rare diseases.
- The model reveals a predictive relationship in which out-of-pocket costs and self-rated health are the strongest signals of medication behavior change.
- Targeting these predictors, future policy could explore a dynamic reimbursement mechanism and strengthen supplementary coverage for rare diseases beyond the basic insurance scheme to significantly improve medication accessibility.

Correspondence: Min Hu, PhD Professor; E-mail: humin@fudan.edu.cn

Presented at ISPOR Real-World Evidence Summit 2025| September 28-30, 2025 | Tokyo, Japan