

Development and Validation of a Machine Learning Model for Predicting Medication Adherence Among Home-Dwelling Elderly Patients: A Retrospective Cross-Sectional Study

Yujie Zhang^{1,2}, Yongli Han², Xuemei Yin³, Yue Tian¹, Mingfen Wu¹

¹Department of Pharmacy, Beijing Tiantan Hospital, Capital Medical University, Beijing, People's Republic of China; ²Department of Pharmacy, Linfen Central Hospital, Linfen, People's Republic of China; ³Department of Pharmacy, Chengshousi Street Community Health Service Center, Beijing, People's Republic of China

Correspondence: Mingfen Wu, Department of Pharmacy, Beijing Tiantan Hospital, Capital Medical University, No. 119 South Fourth Ring West Road, Fengtai District, Beijing, 100070, People's Republic of China, Tel +86-010-59975444, Email wmfing@163.com

Background: Elderly chronic disease management is often complicated by multimorbidity and the need for lifelong polypharmacy, yet poor adherence severely impedes progress. Existing studies mainly focus on exploring influencing factors of medication adherence, while most predictive models lack model interpretability and rarely integrate psychological and social support factors for community elderly populations. Predictive models identifying risk of adherence could enable proactive intervention.

Objective: To develop an interpretable machine learning prediction model that fills the above research gap to predict the medication adherence of elderly patients with chronic diseases in China.

Methods: From January to December 2024, data were collected from chronic disease patients aged 60 years and older receiving home-based medication therapy through face-to-face interviews conducted by pharmacists. Variables included demographic information, comorbidities, chronic diseases and medications information, medication adherence, self-efficacy in rational drug use, medication beliefs, social support, and medication literacy. The dataset was randomly divided into a training set and a test set at a 7:3 ratio. Multivariate logistic regression analysis was performed on all data, and predictors were selected from the training set via the Least Absolute Shrinkage and Selection Operator (LASSO). Six machine learning algorithms were applied in R software to develop predictive models using the training set, and their performance was compared on the test set. The Shapley Additive Explanations (SHAP) approach was used to interpret the optimal model.

Results: A total of 1722 patients were included in the statistical analysis. The gradient boosting machine (GBM) exhibited the best predictive performance among the six models (AUC = 0.811, 95% CI 0.774–0.840), with its core predictors being self-efficacy in rational drug use, medication practice, concern beliefs, and availability of social support. Through SHAP analysis, the interpretability of the model was significantly enhanced, providing a clear decision-making basis for clinicians.

Conclusion: We constructed a prediction model for home medication adherence in elderly patients with chronic diseases, which incorporates important social and psychological factors affecting patients' adherence and provides robust evidence for developing targeted interventions.

Keywords: elderly, chronic diseases, medication adherence, machine learning, prediction model

Introduction

With global population aging, chronic diseases (eg., hypertension, diabetes, coronary heart disease) pose a major public health challenge.^{1,2} In China, accelerating population aging has led to a continuous rise in the number of chronic disease patients. Multimorbidity and polypharmacy are becoming increasingly prevalent, making chronic diseases the primary threat to residents' health, accounting for over 80% of all deaths.^{1,3} While lifelong medication is essential, poor



adherence affects approximately 50% of patients, leading to impaired disease control, higher hospitalization rates, and increased medical costs.⁴ Conversely, improved adherence correlates with better metabolic control, quality of life, reduced complications, and lower hospitalization rates; indeed, each 10% increase in adherence is estimated to save \$450 per patient annually.^{5–9}

Early prediction and targeted home pharmaceutical care are effective strategies to improve adherence.¹⁰ However, given China's 300 million chronic disease patients and the limited resources of home pharmacists, universal interventions are costly, lack targeting, and result in a waste of scarce medical resources. Moreover, most previous studies relied on post-hoc evaluations, missing the window for proactive intervention. Therefore, early identification of high-risk patients is crucial for prioritizing resources and implementing timely care.

Psychosocial factors are modifiable predictors of medication adherence and can be targeted through interventions. Beyond physiological characteristics, social support, self-efficacy, and medication literacy significantly influence adherence behaviors.^{11–16} Pharmacist interventions have been shown to improve these factors and subsequent adherence.¹³

In the existing literature, multiple studies have investigated various factors influencing medication adherence in older adults, including clinical characteristics, demographic attributes, and patients' individualized attitudes toward drug therapy.^{17–19} These studies have laid an important foundation for understanding medication adherence, yet they still have notable limitations: most studies only employed univariate analysis or traditional statistical models, lacking interpretable predictive tools that integrate psychosocial factors, which are key determinants of adherence behavior.

Machine learning (ML) offers a cost-effective approach to developing accurate predictive models.¹¹ Yet, existing domestic models often suffer from small sample sizes, single-center limitations, and insufficient consideration of psychosocial factors in Chinese elderly populations. To overcome these limitations, this study aimed to construct and validate multiple machine learning models based on readily available clinical data. By integrating psychosocial factors including self-efficacy and social support into the predictive framework, the models were applied to predict medication adherence among community-dwelling older adults with chronic diseases in China. Furthermore, the optimal model was selected, and Shapley Additive Explanations (SHAP) were adopted to interpret its prediction outputs.

Materials and Methods

Study Design and Ethics

This study was part of the National Natural Science Foundation of China (No.72404196). A cross-sectional study was conducted in central China, with the survey carried out from January 1 to December 31, 2025. Using a simple sampling method, the sample size was determined based on a 90% power analysis. A total of 1880 patients participated in the questionnaire survey. Prior to the study, informed consent was obtained from all participants. This survey was conducted via face-to-face interviews, each lasting approximately 30 minutes. Given the large number of questions included in the questionnaire, pharmacists communicated with patients using an electronic questionnaire during the interviews and were responsible for recording the patients' responses, so as to avoid respondent fatigue or incomplete answers. The overall length of the questionnaire was manageable for participants and did not have a negative impact on the completion rate. The study was approved by the Ethics Committee of Beijing Tiantan Hospital, Capital Medical University (Approval No. KY2024-408-02).

Data Sources and Participants

This study used a non-experimental, comparative, and predictive research design. The study population comprised elderly patients with chronic diseases receiving long-term medication at home. The objectives were to identify key psychosocial and clinical factors associated with medication adherence and to develop an optimal predictive model for medication adherence.

Participants were recruited from 25 communities across five cities: Beijing, Guangzhou, Sanming, Zhengzhou, and Kunming, between January 1 and December 31, 2025.

Inclusion criteria: Participants were eligible if they met the following criteria: (1) age ≥ 60 years; (2) diagnosed with at least one of four chronic conditions: hypertension, type 2 diabetes mellitus, coronary heart disease, or stroke; (3)

receiving long-term pharmacotherapy; (4) fully informed of the study objectives, willing to participate, and able to provide written informed consent.

Exclusion criteria: Patients were excluded if they had psychiatric disorders, malignant tumors, significant hearing impairment, or communication difficulties.

Recruitment and data collection: Convenience sampling was employed to recruit patients with chronic diseases from various communities. Pharmacists involved in the study completed two rounds of standardised training prior to data collection. Data were collected through structured face-to-face interviews: pharmacists administered the questionnaires verbally, recorded patients' responses, and subsequently submitted the completed forms.

Questionnaire Content

The survey comprised three major sections: (1) Clinical factors: type of chronic disease (hypertension, diabetes, coronary heart disease, stroke and others), number of chronic conditions, disease duration, number of medications used, history of adverse drug reactions, and achievement of treatment targets. (2) Demographic factors: age, sex, marital status, smoking and alcohol use, living arrangement, employment status, educational level, monthly income, and method of medical payment. (3) Psychosocial factors: medication-related self-efficacy, medication beliefs, social support, and medication literacy.

Instruments

Medication adherence: Measured using the sinicized *Adherence to Refills and Medications Scale (ARMS)*.²⁰ Sinicized by two-way translation-back-translation, it is a self-reported scale. It showed good reliability (Cronbach's $\alpha = 0.814$) and content validity (verified by experts).²⁰ This scale includes 12 items rated on a 4-point Likert scale (1 = never, 2 = sometimes, 3 = often, 4 = always). Lower total scores indicate better adherence. A score <16 was classified as good adherence, whereas a score ≥ 16 indicated poor adherence.

Self-efficacy: Assessed using the sinicized *Self-efficacy for Appropriate Medication Use Scale (SEAMS)*.²¹ Sinicized by the same method as ARMS, it is a self-reported scale. It had good reliability (Cronbach's $\alpha = 0.89$).²¹ The scale consists of 13 items, each rated on a 10-point Likert scale (1 = not at all confident, 10 = completely confident). Higher scores indicate greater self-efficacy in managing medication appropriately.

Social support: Evaluated with the sinicized *Social Support Rating Scale (SSRS)*.²² A mature self-reported scale sinicized by standard translation-back-translation, it had good reliability (Cronbach's $\alpha = 0.921$).²² The scale comprises 10 items across three dimensions: objective support (3 items), subjective support (4 items), and utilization of support (3 items). Items 1–4 and 8–10 are scored on a 4-point scale, whereas items 5–7 are rated according to the number of support sources reported. Total scores are classified as low (≤ 22), moderate (23–44), or high (45–66), with higher scores indicating stronger social support.

Medication Beliefs: Assessed with the sinicized *10-item Beliefs about Medicines Questionnaire (BMQ)*.²³ This scale has shown good reliability in previous studies, with a reported Cronbach's alpha coefficient of 0.738.²³ The instrument comprises two subscales: Necessity (5 items) and Concerns (5 items). Each item is rated on a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Subscale scores range from 5 to 25, with higher scores indicating stronger beliefs in the respective dimension. An overall medication-belief score was derived by subtracting the Concerns score from the Necessity score (range: -20 to $+20$). Negative values indicate negative beliefs about medication, whereas higher (more positive) scores indicate stronger positive beliefs.

Medication literacy: The questionnaire was adapted from the *Chinese Medication Literacy KAP scale (KAP)*, with further testing for reliability and validity.²⁴ The adapted scale demonstrated excellent psychometric properties: Cronbach's $\alpha = 0.918$ (>0.80 , indicating high internal consistency), Kaiser–Meyer–Olkin (KMO) value = 0.938 (>0.80), and Bartlett's test of sphericity was highly significant ($p < 0.001$), confirming robust construct validity. The final KAP-based medication literacy scale consists of 41 items across three domains: medication-related knowledge, attitudes, and practices. Higher total scores represent better medication literacy.

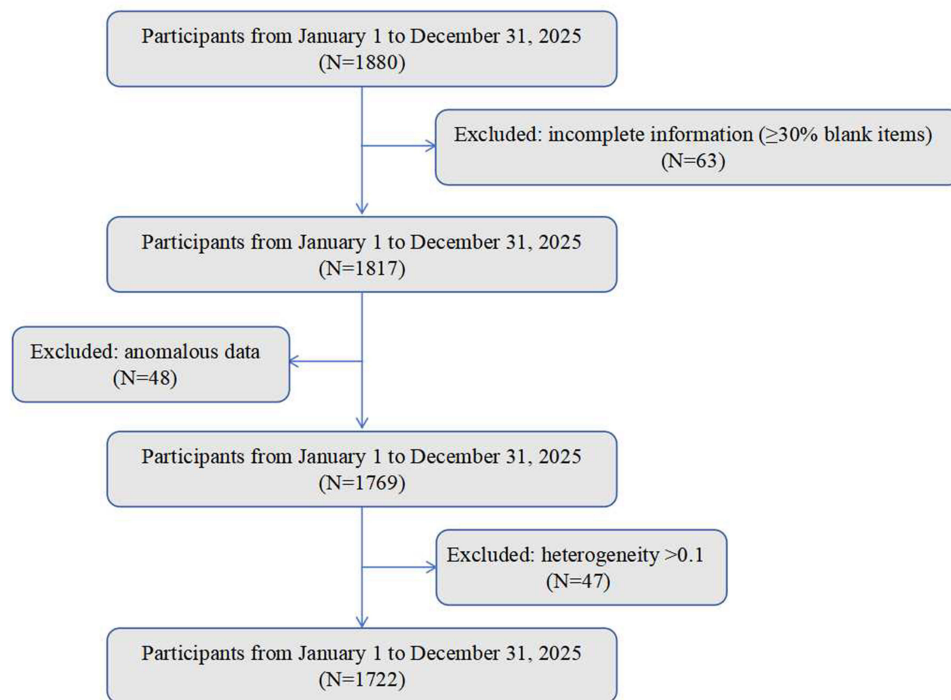


Figure 1 Flowchart of the data preprocessing in the study.

Selection of Predictor Variables

The workflow of data preprocessing is shown in [Figure 1](#). Ultimately, 1722 valid questionnaires were included in the statistical analysis. The dataset comprising 1722 samples was randomly split into a training set (70%) and a test set (30%).

Optimal predictors were identified using the training set through a two-step process. First, multivariate logistic regression analysis was first performed on the full dataset to screen for independent predictors of medication adherence; Second, the Least Absolute Shrinkage and Selection Operator (LASSO) method was then applied exclusively to the training set to further validate and refine the optimal predictor variables. Finally, the four variables consistently identified by both methods were used to construct the prediction model, and the model's discriminative ability, calibration, and clinical utility were evaluated on the test set.

Machine Learning Modeling and Evaluation

The preselected features were input into six distinct machine learning (ML) models to ensure comprehensive and robust evaluation of predictive performance: logistic regression (LR), support vector machine (SVM), decision tree (DT), light gradient boosting machine (LGM), gradient boosting machine (GBM), and extreme gradient boosting (XGB). For each model, Bayesian optimization was used to select hyperparameters, with the objective of maximizing the area under the receiver operating characteristic curve (AUROC) on the training set to ensure optimal performance, and to enable valid comparison of predictive performance on the test set.

All models underwent five-fold cross-validation to enhance robustness and reliability. The predictive performance of each model was evaluated on the test set by plotting receiver operating characteristic (ROC) curves. Additionally, decision curve analysis (DCA) and calibration curve analysis were conducted to assess the clinical utility and prediction accuracy of the models, respectively.

The optimal model was selected based on AUROC values from the test set comparisons. Subsequently, SHAP method was applied to interpret how individual features influence predictions of medication adherence within the optimal model. As a model-agnostic interpretation tool, SHAP quantifies the contribution of each feature to generate both global (model-

level) and local (individual-level prediction) explanations, thereby improving the model's interpretability and clinical applicability.^{25,26}

The training set was used for model development, and the test set for performance evaluation. The primary metric for assessing discriminative ability was the AUROC on the test set. AUROC is a gold-standard method for evaluating a model's ranking ability, as it provides a threshold-independent measure of inherent discriminative power.²⁷ Secondary metrics included sensitivity, specificity, accuracy, recall, F1-score, positive predictive value (PPV), and negative predictive value (NPV) to comprehensively evaluate model performance. Calibration curves were plotted to assess prediction accuracy, and DCA was performed to quantify clinical utility.

Statistical Analysis

All statistical analyses were conducted using R software (version 4.5.0). Sparse data issues were addressed by merging categorical variables where necessary. Since some continuous variables violated the normality assumption (verified via the Shapiro–Wilk test), non-parametric tests were used for between-group comparisons: the Mann–Whitney *U*-test was applied for continuous variables, and the chi-square test (or Fisher's exact test, as appropriate based on data distribution) was used for categorical variables. All tests were two-tailed, and a *P*-value < 0.05 was considered statistically significant.

Results

Baseline Characteristics

Among the 1722 elderly patients, 799 patients showed good medication adherence. The mean age of the participants was 69.72 ± 7.33 years, with 53.3% female and 83.4% married. With regard to educational level, 29.8% had primary education, 63.2% had secondary education, and 7.0% had higher education. A total of 88.9% of patients received daily living care, and approximately 92.5% were unemployed. In terms of economic status, 57.4% were middle-income, 39.2% were low-income, and 3.4% were high-income. The primary type of medical insurance was the basic medical insurance for urban employees (96.1%). The duration of chronic disease exceeded 10 years in 43.3% of patients, and 77.4% of patients reported no adverse drug reactions (ADR).

Significant differences were identified between the good and poor adherence group in the following variables: BMI, number of medications, necessity beliefs, concern beliefs, self-efficacy, subjective support, availability of support, medication knowledge, medication attitude, medication practice, educational level, smoking status, monthly income (CNY), and history of ADR. The demographic and clinical characteristics of patients with good and poor medication adherence are summarized in [Table 1](#).

Selection of Predictor Variables

Variables that were statistically significant and clinically meaningful in the univariate analysis were first included in the multivariate unconditional logistic regression model. The results showed that several variables were independently associated with the outcome ($P < 0.05$), as detailed in [Table 2](#).

Subsequently, the LASSO algorithm was applied for further feature selection ([Figure 2](#)), which identified four key variables as significant predictors of medication adherence (good vs. poor): Concern Beliefs, Self-efficacy, Availability of Support, and Medication Practice.

To verify the consistency of the selection results, we compared the variables identified by multivariate logistic regression and LASSO regression. This comparison revealed that all four core variables screened by LASSO were included in the independent predictors derived from multivariate logistic regression, indicating high consistency between the two methods.

Based on these four core variables jointly identified by the two methods, a final prediction model was constructed, and the model's discriminative ability, calibration, and clinical practical value were further evaluated in the test set.

Table 1 Baseline Characteristics of All Included Patients Stratified by Medication Adherence Status (n=1722)

Variables	Total (N=1722)	Good Adherence (N=799)	Poor Adherence (N=923)	p
Age	69.72±7.33	69.86±7.17	69.61±7.47	0.478
Height	1.65±0.37	1.66±0.53	1.63±0.08	0.214
Weight	65.90±10.62	65.47±10.71	66.28±10.52	0.116
BMI	24.60±3.45	24.38±3.39	24.79±3.49	0.014
Number_of_chronic_diseases	2.94±29.44	3.60±42.75	2.37±5.93	0.387
Number_of_medications	3.57±2.52	3.41±2.47	3.70±2.56	0.019
Necessity_Beliefs	19.06±4.41	19.90±4.61	18.34±4.09	<0.001
Concern_Beliefs	14.36±4.46	13.39±4.95	15.20±3.79	<0.001
Self_efficacy_score	27.10±7.66	30.70±7.39	23.98±6.41	<0.001
Objective_support	8.31±2.51	8.38±2.44	8.24±2.56	0.272
Subjective_support	23.12±5.26	24.29±5.27	22.10±5.04	<0.001
Availability_of_support	7.61±2.38	8.19±2.54	7.11±2.11	<0.001
Medication_knowledge	92.12±19.56	97.45±20.08	87.50±17.86	<0.001
Medication_attitude	33.60±6.78	35.46±7.06	32.00±6.09	<0.001
Medication_practice	68.81±10.72	72.27±10.31	65.81±10.16	<0.001
Gender, n (%)				0.807
Male	805 (46.7%)	371 (46.4%)	434 (47.0%)	
Female	917 (53.3%)	428 (53.6%)	489 (53.0%)	
Marital_status, n (%)				0.492
Married	1436 (83.4%)	661 (82.7%)	775 (84.0%)	
Widowed/Single/Divorced	286 (16.6%)	138 (17.3%)	148 (16.0%)	
Educational_level, n (%)				<0.001
Basic_education	513 (29.8%)	198 (24.8%)	315 (34.1%)	
Higher_education	120 (7.0%)	56 (7.0%)	64 (6.9%)	
Intermediate_education	1089 (63.2%)	545 (68.2%)	544 (58.9%)	
Smoking_or_not, n (%)				0.037
No_smoking	1245 (72.3%)	597 (74.7%)	648 (70.2%)	
Smoking_history	477 (27.7%)	202 (25.3%)	275 (29.8%)	
Drinking_or_not, n (%)				0.441
Drinking_history	494 (28.7%)	222 (27.8%)	272 (29.5%)	
No_drinking	1228 (71.3%)	577 (72.2%)	651 (70.5%)	
Living_arrangement, n (%)				0.530
No_care	192 (11.1%)	85 (10.6%)	107 (11.6%)	
With_care	1530 (88.9%)	714 (89.4%)	816 (88.4%)	
Employment_status, n (%)				0.600
No_employment	1593 (92.5%)	742 (92.9%)	851 (92.2%)	
Still_employed	129 (7.5%)	57 (7.1%)	72 (7.8%)	
Monthly_income_CNY, n (%)				<0.001
High_income	59 (3.4%)	38 (4.8%)	21 (2.3%)	
Low_income	675 (39.2%)	282 (35.3%)	393 (42.6%)	
Middle_income	988 (57.4%)	479 (59.9%)	509 (55.1%)	
Reimbursement_type, n (%)				0.065
Basic_medical_insurance	1654 (96.1%)	760 (95.1%)	894 (96.9%)	
Non_basic_reimbursement	68 (3.9%)	39 (4.9%)	29 (3.1%)	
Course_of_chronic_disease, n(%)				0.272
Long_course	745 (43.3%)	362 (45.3%)	383 (41.5%)	
Medium_long_course	463 (26.9%)	205 (25.7%)	258 (28.0%)	
Short_course	514 (29.8%)	232 (29.0%)	282 (30.6%)	
History_of_ADR, n (%)				<0.001
No_ADR	1333 (77.4%)	664 (83.1%)	669 (72.5%)	
Yes_ADR	389 (22.6%)	135 (16.9%)	254 (27.5%)	

Abbreviations: CNY, Chinese Yuan (¥); p, level of significance; ADR, adverse drug reactions; BMI, Body Mass In.

Table 2 Binary Logistic Regression Analysis of Factors Influencing Medication Adherence

Variables	β	SE	OR (95% CI)	p
(Intercept)	5.823	0.799	338.078(70.592–1619.106)	0.000
Educational_level				
Basic_education			Reference	
Higher_education	-0.301	0.341	0.740(0.379–1.445)	0.377
Intermediate_education	-0.285	0.183	0.752(0.525–1.077)	0.120
Smoking_or_not				
No_smoking			Reference	
Smoking_history	0.411	0.154	1.508(1.114–2.041)	0.008
Monthly_income_CNY				
High_income			Reference	
Low_income	0.161	0.447	1.174(0.489–2.819)	0.719
Middle_income	0.299	0.422	1.349(0.590–3.082)	0.478
History_of_ADR				
No_ADR			Reference	
Yes_ADR	0.209	0.168	1.232(0.887–1.712)	0.213
Necessity_Beliefs	-0.023	0.017	0.977(0.945–1.011)	0.178
Concern_Beliefs	0.066	0.016	1.069(1.035–1.103)	0.000
Self_efficacy_score	-0.104	0.011	0.901(0.881–0.922)	0.000
Subjective_support	-0.018	0.015	0.982(0.954–1.011)	0.225
Availability_of_support	-0.099	0.033	0.906(0.850–0.966)	0.003
Medication_knowledge	-0.003	0.006	0.997(0.985–1.009)	0.622
Medication_attitude	0.013	0.017	1.013(0.980–1.047)	0.436
Medication_practice	-0.037	0.008	0.964(0.949–0.978)	0.000

Notes: All models used binary logistic regression with AMRS as the dependent variable (binary outcome).

Abbreviations: β , regression coefficient; SE, standard error; OR, odds ratio; 95% CI, calculated by the Wald method; p, level of significance; CNY, Chinese Yuan (¥); ADR, adverse drug reactions; BMI, Body Mass Index.

Model Development and Validation

For model development and validation, the 1722 eligible participants were randomly divided into a training set (n=1205, 70%) and a test set (n=517, 30%). All final model predictors and medication adherence status (good/poor) were balanced between the training set and the test set, as presented in Table 3.

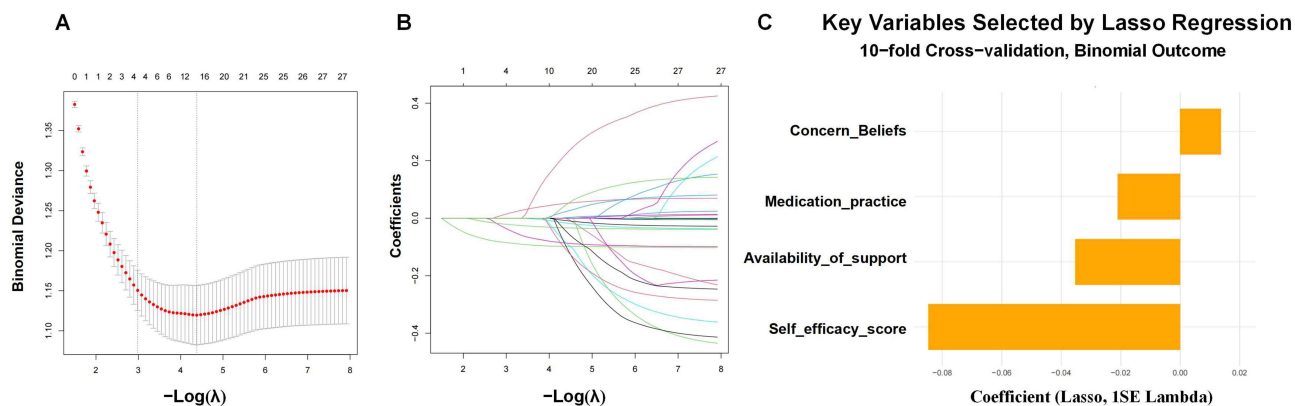


Figure 2 Feature selection based on the LASSO algorithm. (A) Selection of the tuning parameter (λ) in the LASSO model via 10-fold cross-validation based on minimum criteria. The optimal λ value of 0.051. (B) The LASSO coefficient profiles of clinical features. (C) The coefficients of LASSO regression analysis.

Table 3 Characteristics of the Predictors in the Training and Testing Sets

Variables	Total (N=1722)	Testdata (N=517)	Triandata (N=1205)	p
Age, Mean±SD	69.722±7.331	69.888±7.348	69.651±7.326	0.540
Height, Mean±SD	1.645±0.369	1.637±0.081	1.649±0.438	0.548
Weight, Mean±SD	65.905±10.615	66.346±10.440	65.716±10.689	0.259
BMI, Mean±SD	24.598±3.450	24.719±3.163	24.546±3.566	0.341
Number_of_chronic_diseases, Mean±SD	2.938±29.437	2.172±1.151	3.267±35.181	0.479
Number_of_medications, Mean±SD	3.566±2.518	3.625±2.297	3.540±2.608	0.523
Necessity_Beliefs, Mean±SD	19.063±4.408	18.946±4.496	19.114±4.371	0.469
Concern_Beliefs, Mean±SD	14.363±4.460	14.186±4.259	14.439±4.543	0.280
Self_efficacy_score, Mean±SD	27.096±7.655	26.584±7.702	27.315±7.628	0.069
Objective_support, Mean±SD	8.307±2.508	8.414±2.529	8.261±2.498	0.245
Subjective_support, Mean±SD	23.120±5.259	22.897±5.268	23.216±5.254	0.250
Availability_of_support, Mean±SD	7.614±2.382	7.588±2.385	7.625±2.382	0.768
Medication_knowledge, Mean±SD	92.117±19.557	91.089±19.031	92.558±19.770	0.153
Medication_attitude, Mean±SD	33.605±6.781	33.168±6.381	33.792±6.940	0.080
Medication_practice, Mean±SD	68.811±10.721	68.491±10.775	68.949±10.699	0.417
Adherence_group, n (%)				0.468
Good	799 (46.4%)	233 (45.1%)	566 (47.0%)	
Poor	923 (53.6%)	284 (54.9%)	639 (53.0%)	
Gender, n (%)				0.381
Female	917 (53.3%)	267 (51.6%)	650 (53.9%)	
Male	805 (46.7%)	250 (48.4%)	555 (46.1%)	
Marital_status, n (%)				0.763
Married	1436 (83.4%)	429 (83.0%)	1007 (83.6%)	
Widowed/Single/Divorced	286 (16.6%)	88 (17.0%)	198 (16.4%)	
Educational_level, n (%)				0.697
Basic_education	513 (29.8%)	161 (31.1%)	352 (29.2%)	
Higher_education	120 (7.0%)	34 (6.6%)	86 (7.1%)	
Intermediate_education	1089 (63.2%)	322 (62.3%)	767 (63.7%)	
Smoking_or_not, n (%)				0.335
No_smoking	1245 (72.3%)	382 (73.9%)	863 (71.6%)	
Smoking_history	477 (27.7%)	135 (26.1%)	342 (28.4%)	
Drinking_or_not, n (%)				0.279
Drinking_history	494 (28.7%)	139 (26.9%)	355 (29.5%)	
No_drinking	1228 (71.3%)	378 (73.1%)	850 (70.5%)	
Living_arrangement, n (%)				0.575
No_care	192 (11.1%)	61 (11.8%)	131 (10.9%)	
With_care	1530 (88.9%)	456 (88.2%)	1074 (89.1%)	
Employment_status, n (%)				0.650
No_employment	1593 (92.5%)	476 (92.1%)	1117 (92.7%)	
Still_employed	129 (7.5%)	41 (7.9%)	88 (7.3%)	
Monthly_income_CNY, n (%)				0.735
High_income	59 (3.4%)	15 (2.9%)	44 (3.7%)	
Low_income	675 (39.2%)	204 (39.5%)	471 (39.1%)	
Middle_income	988 (57.4%)	298 (57.6%)	690 (57.3%)	
Reimbursement_type, n (%)				0.083
Basic_medical_insurance	1654 (96.1%)	503 (97.3%)	1151 (95.5%)	
Non_basic_reimbursement	68 (3.9%)	14 (2.7%)	54 (4.5%)	
Course_of_chronic_disease, n (%)				0.103
Long_course	745 (43.3%)	208 (40.2%)	537 (44.6%)	
Medium_long_course	463 (26.9%)	156 (30.2%)	307 (25.5%)	
Short_course	514 (29.8%)	153 (29.6%)	361 (30.0%)	

(Continued)

Table 3 (Continued).

Variables	Total (N=1722)	Testdata (N=517)	Triandata (N=1205)	p
History_of_ADR, n (%)				0.199
No_ADR	1333 (77.4%)	390 (75.4%)	943 (78.3%)	
Yes_ADR	389 (22.6%)	127 (24.6%)	262 (21.7%)	

Note: Data are presented as mean \pm standard deviation or number (%).

Abbreviations: CNY, Chinese Yuan (¥); p, level of significance; ADR, adverse drug reactions; BMI, Body Mass Index.

The discriminative ability of the six ML models in the training set was evaluated using AUROC, with results shown in [Figure 3A](#). Among the six models, the XGB model exhibited the highest discrimination, with an AUC of 0.850 (95% CI: 0.829–0.871). The LGB model followed closely (AUC=0.845, 95% CI: 0.824–0.867), while the GBM model had an AUC of 0.843 (95% CI: 0.821–0.865). The SVM model had an AUC of 0.813 (95% CI: 0.789–0.837), and the LR and DT models had the lowest AUC values, at 0.792 (95% CI: 0.767–0.817) and 0.769 (95% CI: 0.743–0.796), respectively.

The AUROC values of the six models in the test set (used for external validation) are shown in [Figure 3B](#). In contrast to the training set results, the GBM model showed the highest discrimination in the test set, with an AUC of 0.811 (95% CI: 0.774–0.840). The XGB model followed closely (AUC=0.806, 95% CI: 0.768–0.843), and the LGB model had an AUC of 0.802 (95% CI: 0.765–0.840). The SVM model had an AUC of 0.792 (95% CI: 0.753–0.831), while the DT and LR models had the lowest AUC values, at 0.767 (95% CI: 0.726–0.808) and 0.763 (95% CI: 0.722–0.804), respectively.

[Table 4](#) presents the detailed prediction performance of the 6 machine learning (ML) models in the test set. The GBM model achieved the highest AUC of 0.811 (95% CI: 0.774–0.840), while the LR model had the lowest AUC of 0.763 (95% CI: 0.722–0.804). Regarding other key performance indicators, the DT model exhibited the best sensitivity (0.877), the LGB model had the best specificity (0.768), and the XGB model showed the highest accuracy (0.743). [Figure 3C](#) shows the calibration curves of all prediction models in the training set, while the calibration curves of the prediction models in the test set are presented in [Figure 3D](#). The LGB model exhibited the best calibration performance in the test set.

DCA was performed to evaluate the clinical benefit of each prediction model, with DCA curves in the training set shown in [Figure 3E](#) and those in the test set presented in [Figure 3F](#). In the test set, the effective threshold range of the DT model was approximately 0–0.76, which was significantly narrower than that of the GBM (0–0.88), XGB (0–0.87), LGB (0–0.87), SVM (0–0.87), and LR (0–0.85) models. Additionally, within the test set, in the threshold range of 0–0.11, the net benefits of all six models were completely consistent (net benefit=0.4992 for all models at a threshold of 0.1), indicating no significant difference in net benefits across models within this range; Within the threshold range of 0.7–0.85, the DCA curves of the GBM (net benefit: 0.1509 at 0.7, 0.1257 at 0.8), XGB (net benefit: 0.1373 at 0.7, 0.1257 at 0.8), LGB (net benefit: 0.1567 at 0.7, 0.1335 at 0.8), and SVM (net benefit: 0.1567 at 0.7, 0.1335 at 0.8) models partially overlapped, with differences in net benefits ranging only from 0.0032 to 0.0194, also indicating no significant differences in net benefits among these models within this range.

Comprehensively considering the AUROC (primary discriminative indicator), sensitivity, and specificity (secondary performance indicators), the gradient boosting machine (GBM) model performed optimally and was thus identified as the best model for predicting medication adherence in community-dwelling elderly individuals.

SHAP-Based Model Interpretation

To intuitively present the selected variables, SHAP was used to illustrate how these features predict medication adherence in community-dwelling elderly individuals in the GBM model ([Figure 4](#)). [Figure 4A](#) presents the four features ranked by their mean absolute SHAP values, where a higher mean absolute SHAP value indicates a greater contribution to the risk of medication adherence. [Figure 4B](#) illustrates the impact values and interpretations of these features, with yellow dots representing high risk and purple dots representing low risk. Self-efficacy Score, Medication Practice, Concern Beliefs, and Availability of Support were associated with a higher risk of medication non-adherence in community-dwelling elderly patients with chronic diseases. In addition to the global SHAP interpretation, the local

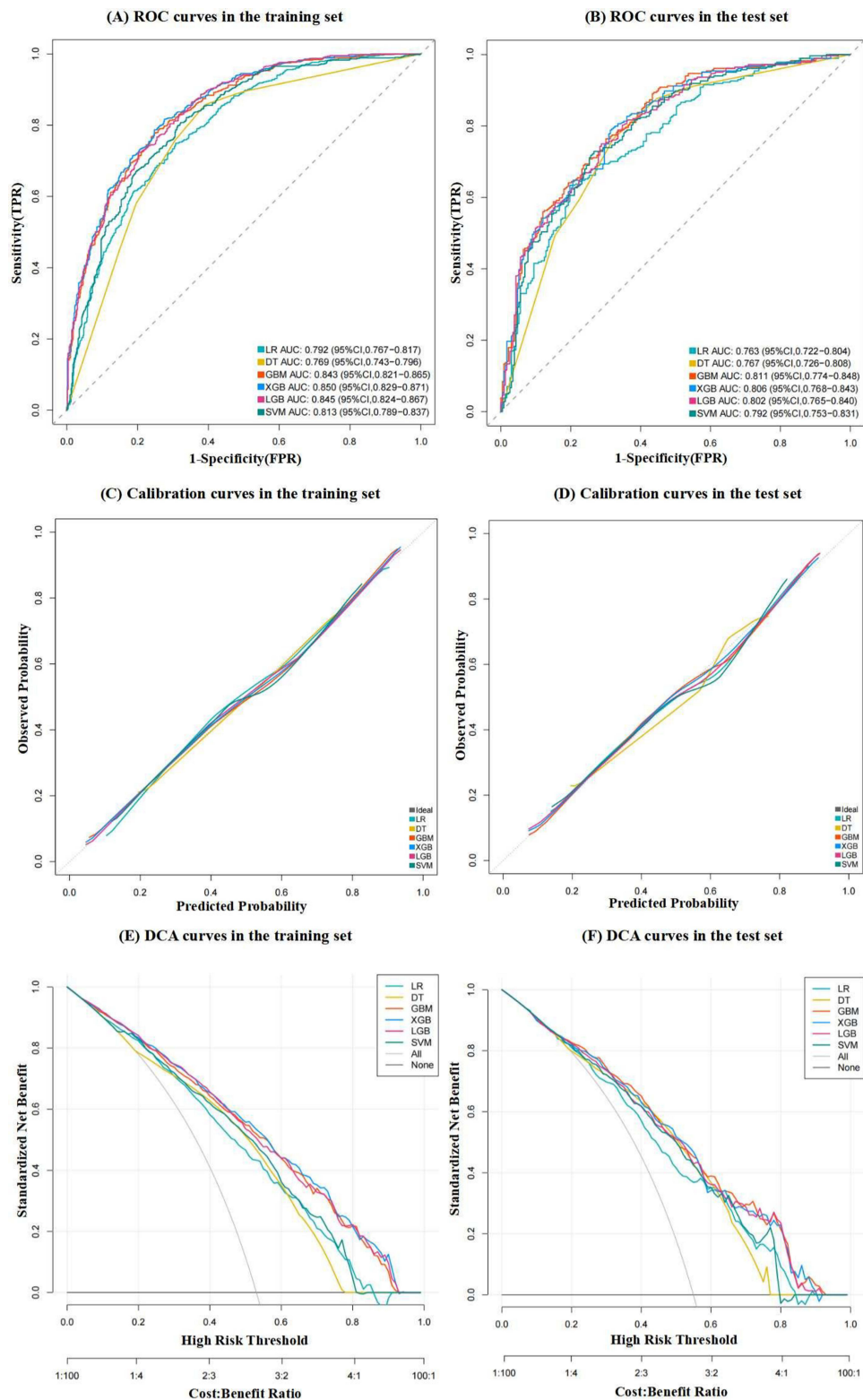


Figure 3 Receiver operating characteristic (ROC) curves, calibration curves, and decision curve analysis (DCA) curves of different machine learning (ML) models in the training and testing sets. **(A)** ROC curves and area under the ROC curve (AUROC) values of different ML prediction models in the training set. **(B)** ROC curves and AUC values of different ML prediction models in the testing set. **(C)** Calibration curves of different ML prediction models in the training set. **(D)** Calibration curves of different ML prediction models in the testing set. **(E)** DCA curves of different ML prediction models in the training dataset. **(F)** DCA curves of different ML prediction models in the testing dataset.

Abbreviations: GBM, gradient boosting machine; LR, logistic regression; SVMs, support vector machines; XGBoost, extreme gradient boosting; LGB, Light Gradient Boosting Machine; DT, Decision Tree.

Table 4 The Performance of 7 Machine Learning Models for Predicting Medication Adherence Among Community-Dwelling Older Adults

Model	AUC (95% CI)	Sensitivity	Specificity	F ₁ -score	Accuracy	PPV	NPV	Recall
Training set								
LR	0.792(0.767–0.817)	0.750	0.693	0.742	0.723	0.734	0.71	0.75
DT	0.769(0.743–0.796)	0.858	0.610	0.779	0.741	0.713	0.791	0.858
GBM	0.843(0.821–0.865)	0.778	0.753	0.779	0.766	0.78	0.75	0.778
XGB	0.850(0.829–0.871)	0.817	0.724	0.793	0.773	0.77	0.778	0.817
LGB	0.845(0.824–0.867)	0.723	0.797	0.760	0.758	0.801	0.718	0.723
SVM	0.813(0.789–0.837)	0.808	0.678	0.772	0.747	0.739	0.757	0.808
Testing set								
LR	0.763(0.722–0.804)	0.704	0.665	0.711	0.687	0.719	0.649	0.704
DT	0.767(0.726–0.808)	0.877	0.554	0.782	0.731	0.705	0.787	0.877
GBM	0.811(0.774–0.848)	0.736	0.704	0.744	0.721	0.752	0.686	0.736
XGB	0.806(0.768–0.843)	0.799	0.674	0.773	0.743	0.749	0.734	0.799
LGB	0.802(0.765–0.840)	0.665	0.768	0.717	0.712	0.778	0.653	0.665
SVM	0.792(0.753–0.831)	0.789	0.639	0.757	0.721	0.727	0.713	0.789

Abbreviations: UPRC, area under the precision recall curve; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; DT, Decision Tree; GBM, gradient boosting machine; XGB, extreme gradient boosting; LGB, LGM; SVM, support vector machine.

interpretability of the GBM model was also verified, which could explain the contribution of each feature to the medication adherence prediction result of individual patients. Figure 5 visualizes how the GBM model predicts medication adherence in community-dwelling elderly patients with chronic diseases; yellow arrows indicate features that increase the risk. The $f(x)$ values inside the arrows quantify the contribution of each feature, and the sum of these values yields the final prediction result of the model, which is represented by the $f(x)$ value outside the arrows.

Discussion

In this study, we aimed to construct and compare multiple ML models to predict medication adherence among community-dwelling elderly patients with chronic diseases. We first identified statistically significant variables using multivariate logistic regression, followed by feature selection via the LASSO regression. Ultimately, four core predictors

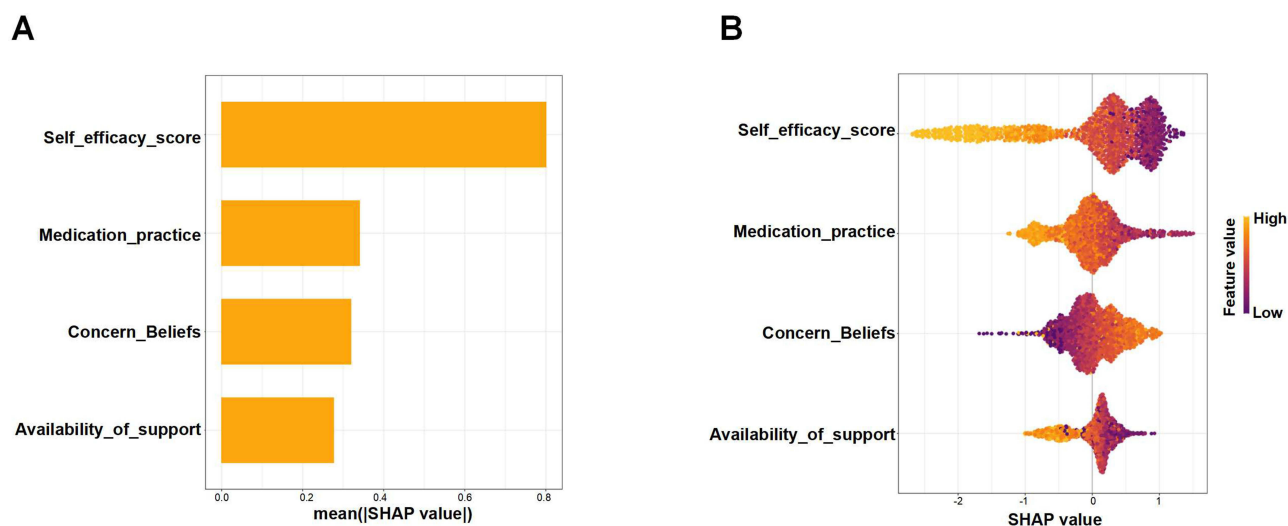


Figure 4 Interpretation of the gradient boosting machine (GBM) model by the Shapley Additive Explanations (SHAP) method. (A) A bar summary of the most important features according to the SHAP values. (B) Summary and explanation of the most influential features. Yellow dots indicate high-risk values, and purple dots indicate low-risk values.

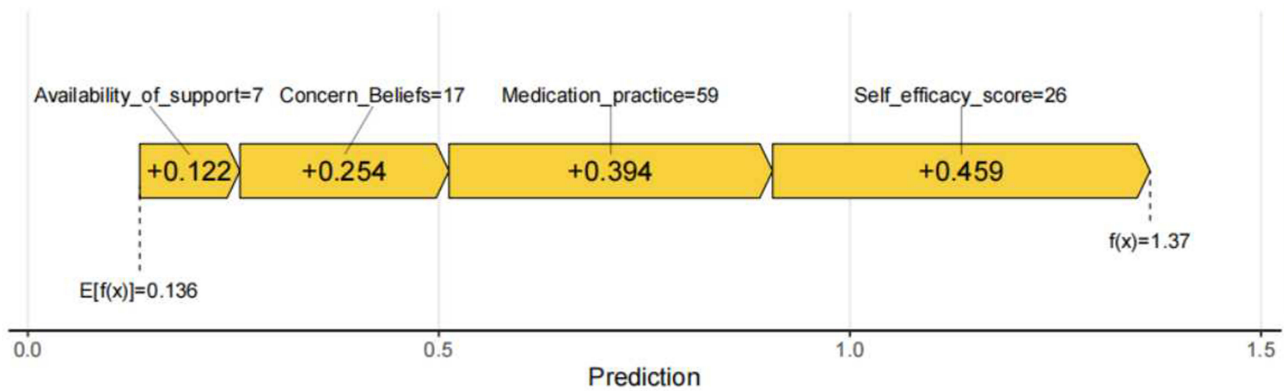


Figure 5 Compositional risk of feature contributions for predicting the occurrence of medication adherence. Purple arrows denote factors that decrease the risk of medication adherence, while yellow arrows indicate features that increase the risk.

were identified: the self-efficacy score, medication practice, concern beliefs, and availability of support. Based on these four variables, we developed and validated six ML prediction models, including GBM, XGB, and LGM. The results demonstrated that the GBM model achieved the highest predictive performance (AUC = 0.811 in the test set) and favorable clinical utility, as evaluated by DCA curves. Furthermore, we quantified the contribution of each feature to the prediction results using SHAP analysis, and found that the self-efficacy score and medication practice were the two most critical factors influencing medication adherence.

Comparison with Prior Work

An increasing number of studies have demonstrated that medication adherence among community-dwelling elderly individuals represents a complex, multifaceted issue influenced by a diverse array of factors, including demographics, psychosocial aspects, clinical characteristics, and medication-related attributes. Previous research has identified numerous risk factors associated with non-adherence, such as patients' age, gender, educational level, attitudes toward medications, beliefs about drug therapy, and social support networks.^{17,18,28,29} However, the majority of these studies remain primarily confined to the identification of risk factors, failing to fully leverage these factors to construct effective predictive models for identifying high-risk populations. Traditional statistical approaches, such as multivariate regression models, exhibit inherent limitations when addressing complex nonlinear relationships within high-dimensional datasets.³⁰ In contrast, machine learning methodologies offer distinct advantages in tackling such intricate problems, enabling the automatic extraction of complex patterns and relationships within the data.^{31–34}

Traditional methods usually employ a single approach (eg., regression models) to select predictors, whereas combining multiple feature selection techniques may yield simplified models with higher generalization ability.³⁵ The key advantage of this study is that we not only compared the performance of multiple machine learning models but also combined traditional multivariate logistic regression with advanced LASSO regression for feature selection, thereby ensuring the simplicity and generalization capability of the final model.³⁶ The four core predictors we screened including self-efficacy score, medication practice, concern beliefs, and availability of support are highly consistent with the conclusions of existing studies, which further verifies the importance of these factors in the prediction of medication adherence.^{37–39}

Interpretation of Core Predictors

Although an increasing number of ML-based clinical prediction models are being developed, most studies lack the interpretability of these models, which limits their clinical understanding and practical adoption. The interpretability of ML predictions requires the urgent attention of researchers, so that clinicians can understand, trust, and ultimately apply these prediction models to guide their clinical practice.^{40–42} This study revealed the specific impact of each predictor on medication adherence through SHAP analysis. Self-efficacy was the most important predictor, and a lower self-efficacy

score was associated with a higher risk of medication non-adherence. This is consistent with the results of numerous studies, indicating that patients' confidence in their ability to perform medication-taking behaviors is crucial for adherence to treatment.⁴³ Patients with high self-efficacy are more confident in coping with complex medication regimens, managing drug side effects, and proactively overcoming difficulties encountered during medication administration.^{14,44}

Medication literacy is the foundation of safe and rational medication use and has also been recognized as an important factor affecting adherence.⁴⁵ Medication practice was the second most important predictor. Medication practice refers to patients' actual medication-taking behaviors, such as whether they take medications on time and in the correct dosage, and whether they adjust medications in accordance with medical advice, which is the most direct external manifestation of medication adherence. A study by Tang et al showed that medication practice scores are associated with medication adherence.⁴⁶ Standardized medication practice can not only better control the disease but also enhance patients' confidence in treatment, forming a positive cycle. The results of this study showed that patients with low medication practice scores had a significantly higher risk of medication non-adherence.

Concern beliefs were also identified as an important predictor. Excessive worries about drug side effects, the long-term safety of medication, or potential drug dependence can lead to hesitation or even refusal to take medication, thereby reducing medication adherence.⁴⁷ This aligns with our findings that higher concern beliefs are associated with poorer medication adherence.

Finally, availability of support was also confirmed to be associated with medication adherence. Support from family, friends, or the community can help patients better understand and implement medication regimens, remind them to take medications on time, thereby improving medication adherence. A study of 259 patients in a cardiology clinic of a university hospital in Turkey found that medication adherence increased positively with the increase of social support in hypertensive patients.⁴⁸ Patients with low support utilization may develop dependence when facing excessive help, which weakens their motivation for self-management, while those with high support utilization may accurately call on resources to form a "reminder–feedback–reinforcement" loop.

Importance of Model Interpretability

Although machine learning (ML) models generally outperform traditional statistical models in terms of predictive performance, their "black-box" nature often limits their clinical acceptance and practical application. To address this issue, we employed the SHAP method to interpret our optimal GBM model. SHAP values quantify the marginal contribution of each input feature to an individual prediction, thereby revealing the model's decision-making process at both the global and local levels. Through SHAP analysis, we not only confirmed the importance ranking of the four core variables but also visually demonstrated the specific direction and extent of each variable's influence on prediction results across different value ranges. This transformed the model's predictions from abstract probability values into interpretable, clinically meaningful explanations that can be understood and trusted by clinicians, thereby significantly enhancing the clinical utility and operability of our model.

Clinical Implications and Future Perspectives

The GBM prediction model constructed in this study has promising clinical application prospects. As a clinical decision support tool, this model can help healthcare providers quickly identify high-risk patients with poor medication adherence, thereby enabling timely implementation of targeted intervention measures, such as strengthening medication education, providing medication reminder services, and optimizing social support. These interventions are expected to improve patients' medication adherence and ultimately enhance treatment outcomes.

In view of the favorable predictive performance of the GBM model (AUC = 0.811), we further formulated targeted and operable communication recommendations for physicians and pharmacists according to the four core influencing factors identified in this study: self-efficacy, medication practice, concern beliefs, and availability of support. Combined with real-world medication counseling experience in community clinical and pharmacy settings, we summarized concise

and patient-oriented communication examples to assist frontline medical staff in conducting individualized health education and adherence interventions for elderly patients, as detailed below:

Self-Efficacy

You are fully capable of managing your daily medication schedule. We can help you develop a regular medication routine. Sticking to standardized medication will effectively control your chronic condition, and you can gradually build stable confidence in long-term medication management.

Medication Practice

Please take your medicine at a fixed time and exact dosage as recommended. Do not skip, reduce or discontinue medication without medical advice. Standard daily medication behavior is essential for stable disease control and long-term health maintenance.

Concern Beliefs About Medicines

All prescribed medications are tailored to your personal physical condition. The clinical benefits of sustained regular medication far outweigh the risk of minor side effects. Unnecessary worries or deliberate refusal to medication should be avoided.

Availability of Support

You may actively rely on family members and community resources for medication reminding and daily care. Making full use of the availability of support around you can greatly help you maintain persistent and good medication adherence.

In the future, we plan to develop this model into a user-friendly online assessment tool or embedded application, which will be integrated into hospital information systems to realize real-time dynamic assessment and early warning of patients' medication adherence risk.

Limitations

This study has several limitations. First, this is a cross-sectional study, and the generalizability of the study results may be limited. Larger-scale, multi-center prospective studies are needed for external validation to confirm the applicability of our findings. In addition, the participants enrolled in this study were aged 60 years and older, which conforms to the age classification standard for older adults in China; however, the sample was recruited from limited regional areas and cannot fully represent the nationwide elderly population aged 65 years and above across all geographical regions of China. Therefore, caution should be exercised when extrapolating the present findings to a broader national population. Second, the data on medication adherence and related variables used in this study mainly relied on patient self-reported questionnaires, which may be subject to recall bias and social desirability bias, thereby affecting the accuracy of the data. In addition, our model only included demographic and psychosocial characteristics. Future studies may consider incorporating more clinical variables and objective measurement indicators to further improve the predictive performance of the model.

Conclusions

In summary, this study successfully constructed and validated a gradient boosting machine (GBM)-based prediction model for medication adherence among community-dwelling elderly patients with chronic diseases. This model can effectively predict the risk of medication non-adherence in patients, with its core predictors being self-efficacy, medication practice, concern beliefs, and availability of support. Through SHAP analysis, the interpretability of the model was significantly enhanced, providing clinicians with a clear basis for clinical decision-making. This model is expected to be widely applied in clinical practice, offering strong support for improving medication management and health outcomes in community-dwelling elderly patients with chronic diseases.

Abbreviations

LASSO, least absolute shrinkage and selection operator; ARMS, adherence to refills and medications scale; SEAMS, self-efficacy for appropriate medication use scale; KAP, knowledge, attitude, and practice; SSRS, social support rating scale; BMQ, beliefs about medicines questionnaire; ChMLM, Chinese medication literacy measure; ML, machine learning; AUPRC, area under the precision recall curve; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; DT, decision tree; GBM, gradient boosting machine; XGB, extreme gradient boosting; LGB, light gradient boosting machine; SVM, support vector machine; CNY, Chinese yuan (¥); ADR, adverse drug reactions; BMI, body mass index; AUROC, area under the receiver operating characteristic curve; DCA, decision curve analysis; SHAP, shapley additive explanations.

Data Sharing Statement

The de-identified data that support the findings of this study are not publicly available due to restrictions imposed by the ethics committee to protect participant privacy. The datasets used and/or analysed available from the corresponding author on reasonable request.

Ethics Approval and Informed Consent

Ethical approval for this study was conducted by the Ethics Committee of Beijing Tiantan Hospital, Capital Medical University (Approval No. KY2024-408-02). The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki.

Consent to Participate

All participants were informed about the purpose of the study, assured of confidentiality, and provided written consent prior to participation. Participation was voluntary, and respondents could withdraw at any time without consequence.

Consent for Publication

Not applicable. This manuscript contains no personal data from any individual. All the authors have agreed to the publication of this manuscript.

Acknowledgments

The authors would like to thank all the participants and colleagues who contributed to this research.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 72404196).

Disclosure

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. *J Am Coll Cardiol.* 2020;76(25):2982–3021. doi:10.1016/j.jacc.2020.11.010

2. Arai H, Ouchi Y, Toba K, et al. Perspectives from medicine and medical care in Japan: aging as the front-runner of super-aged societies. *Geriatrics Gerontol Int.* 2015;15(6):673–687.
3. Bai J, Cui J, Shi F, Yu C. Global epidemiological patterns in the burden of main non-communicable diseases, 1990-2019: relationships with socio-demographic index. *Int J Public Health.* 2023;68:1605502. doi:10.3389/ijph.2023.1605502
4. Brown MT, Bussell JK. Medication adherence: WHO cares? *Mayo Clin Proc.* 2011;86(4):304–314.
5. Lin LK, Sun Y, Heng BH, Chew DEK, Chong PN. Medication adherence and glycemic control among newly diagnosed diabetes patients. *BMJ Open Diabetes Res Care.* 2017;5(1):e000429. doi:10.1136/bmjdr-2017-000429
6. Khayyat SM, Mohamed MMA, Khayyat SMS, et al. Association between medication adherence and quality of life of patients with diabetes and hypertension attending primary care clinics: a cross-sectional survey. *Qual Life Res.* 2019;28(4):1053–1061. doi:10.1007/s11136-018-2060-8
7. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. *Med Care.* 2005;43(6):521–530. doi:10.1097/01.mlr.0000163641.86870.af
8. Evans M, Engberg S, Faurby M, Fernandes JDDR, Hudson P, Polonsky W. Adherence to and persistence with antidiabetic medications and associations with clinical and economic outcomes in people with type 2 diabetes mellitus: a systematic literature review. *Diabetes Obes Metab.* 2022;24(3):377–390. doi:10.1111/dom.14603
9. Kennedy-Martin T, Boye KS, Peng X. Cost of medication adherence and persistence in type 2 diabetes mellitus: a literature review. *Patient Prefer Adherence.* 2017;11:1103–1117. doi:10.2147/PPA.S136639
10. Wu M, Xu X, Zhao R, Bai X, Zhu B, Zhao Z. Effect of pharmacist-led interventions on medication adherence and glycemic control in type 2 diabetic patients: a study from the chinese population. *Patient Prefer Adherence.* 2023;17:119–129. doi:10.2147/PPA.S394201
11. Berry KN, Daniels N, Ladin K. Should lack of social support prevent access to organ transplantation? *Am J Bioeth.* 2019;19(11):13–24. doi:10.1080/15265161.2019.1665728
12. Allen J, Markovitz J, Jacobs DR, Knox SS. Social support and health behavior in hostile black and white men and women in CARDIA. Coronary artery risk development in young adults. *Psychosom Med.* 2001;63(4):609–618. doi:10.1097/00006842-200107000-00014
13. Criswell TJ, Weber CA, Xu Y, Carter BL. Effect of self-efficacy and social support on adherence to antihypertensive drugs. *Pharmacotherapy.* 2010;30(5):432–441.
14. Shen Z, Shi S, Ding S, Zhong Z. Mediating effect of self-efficacy on the relationship between medication literacy and medication adherence among patients with hypertension. *Front Pharmacol.* 2020;11:569092. doi:10.3389/fphar.2020.569092
15. Neiva Pantuzza LL, Nascimento ED, Crepalde-Ribeiro K, et al. Medication literacy: a conceptual model. *Res Social Adm Pharm.* 2022;18(4):2675–2682. doi:10.1016/j.sapharm.2021.06.003
16. Zhu L, Liu Y, Yang F, Yu S, Fu P, Yuan H. Prevalence, associated factors and clinical implications of medication literacy linked to frailty in hemodialysis patients in China: a cross-sectional study. *BMC Nephrol.* 2023;24(1):307. doi:10.1186/s12882-023-03346-4
17. Foley L, Larkin J, Lombard-Vance R, et al. Prevalence and predictors of medication non-adherence among people living with multimorbidity: a systematic review and meta-analysis. *BMJ Open.* 2021;11(9):e044987. doi:10.1136/bmjopen-2020-044987
18. Wilder ME, Kulie P, Jensen C, et al. The impact of social determinants of health on medication adherence: a systematic review and meta-analysis. *J Gen Intern Med.* 2021;36(5):1359–1370. doi:10.1007/s11606-020-06447-0
19. Rhudy C, Johnson J, Perry C, et al. Machine learning approaches to predicting medication nonadherence: a scoping review. *Int J Med Inform.* 2025;204:106082. doi:10.1016/j.ijmedinf.2025.106082
20. Kripalani S, Risser J, Gatti ME, et al. Development and evaluation of the Adherence to Refills and Medications Scale (ARMS) among low-literacy patients with chronic disease. *Value Health.* 2009;12(1):118–123. doi:10.1111/j.1524-4733.2008.00400.x
21. Risser J, Jacobson TA, Kripalani S. Development and psychometric evaluation of the Self-efficacy for Appropriate Medication Use Scale (SEAMS) in low-literacy patients with chronic disease. *J Nurs.* 2007;15(3):203–219.
22. Xiao SY. The theoretical basis and research application of “Social support rating scale”. *J Clin Psychiatry.* 1994;02:98–100.
23. Horne R, Weinman J, Hankins M. The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. *Psychol Health.* 1999;14(1):1–24. doi:10.1080/08870449908407311
24. Zhou B, Zhao Z, Wu M, et al. A medication literacy instrument for older chronic disease patients in China: development and validation. *BMC Geriatr.* 2026. doi:10.1186/s12877-026-07171-w
25. Liu C, Zhang K, Yang X, et al. Development and validation of an explainable machine learning model for predicting myocardial injury after noncardiac surgery in two centers in China: retrospective study. *JMIR Aging.* 2024;7:e54872. doi:10.2196/54872
26. Wang Z, Gu Y, Huang L, et al. Construction of machine learning diagnostic models for cardiovascular pan-disease based on blood routine and biochemical detection data. *Cardiovasc Diabetol.* 2024;23(1):351. doi:10.1186/s12933-024-02439-0
27. Adeli V, Korhani N, Sabo A, et al. Ambient monitoring of gait and machine learning models for dynamic and short-term falls risk assessment in people with dementia. *IEEE J Biomed Health Inform.* 2023;27(7):3599–3609. doi:10.1109/JBHI.2023.3267039
28. Davis DP, Jandrisevits MD, Iles S, Weber TR, Gallo LC. Demographic, socioeconomic, and psychological factors related to medication non-adherence among emergency department patients. *J Emerg Med.* 2012;43(5):773–785. doi:10.1016/j.jemermed.2009.04.008
29. Huang YM, Shiyabola OO, Smith PD. Association of health literacy and medication self-efficacy with medication adherence and diabetes control. *Patient Prefer Adherence.* 2018;12:793–802. doi:10.2147/PPA.S153312
30. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods.* 2018;15(4):233–234. doi:10.1038/nmeth.4642
31. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15(4):361–387. doi:10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4
32. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *BMJ.* 2009;338:b604. doi:10.1136/bmj.b604
33. Battineni G, Sagaro GG, Chinatalapudi N, Amenta F. Applications of machine learning predictive models in the chronic disease diagnosis. *J Pers Med.* 2020;10(2):21. doi:10.3390/jpm10020021
34. Eloranta S, Boman M. Predictive models for clinical decision making: deep dives in practical machine learning. *J Intern Med.* 2022;292(2):278–295. doi:10.1111/joim.13483
35. Wu X, Tang F, Li H, et al. Development and validation of a nomogram model for medication non-adherence in patients with chronic kidney disease. *J Psychosom Res.* 2023;171:111385. doi:10.1016/j.jpsychores.2023.111385

36. Lin B, Feng S, Liu J, Li K, Shi G, Zhong X. Using an interactive web application to identify pre-exposure prophylaxis adherence among men who have sex with men. *Int J Clin Health Psychol.* 2024;24(3):100490. doi:10.1016/j.ijchp.2024.100490
37. Al-Qerem W, Jarab A, Eberhardt J, et al. Medication adherence among Jordanian adults with chronic conditions: a combined analysis using regression and machine learning. *Ann Med.* 2025;57(1):2548979. doi:10.1080/07853890.2025.2548979
38. Wang L, Fan R, Zhang C, et al. Applying machine learning models to predict medication nonadherence in crohn's disease maintenance therapy. *Patient Prefer Adherence.* 2020;14:917–926. doi:10.2147/PPA.S253732
39. Liu L, Yu Z, Chen H, et al. Imatinib adherence prediction using machine learning approach in patients with gastrointestinal stromal tumor. *Cancer.* 2025;131(1):e35548. doi:10.1002/cncr.35548
40. Song Y, Yuan Q, Liu H, Gu K, Liu Y. Machine learning algorithms to predict mild cognitive impairment in older adults in China: a cross-sectional study. *J Affect Disord.* 2025;368:117–126. doi:10.1016/j.jad.2024.09.059
41. Kang CW, Yan ZK, Tian JL, Pu XB, Wu LX. Constructing a fall risk prediction model for hospitalized patients using machine learning. *BMC Public Health.* 2025;25(1):242. doi:10.1186/s12889-025-21284-8
42. Shen L, Jin Y, Pan A, et al. Machine learning-based predictive models for perioperative major adverse cardiovascular events in patients with stable coronary artery disease undergoing noncardiac surgery. *Comput Methods Programs Biomed.* 2025;260:108561. doi:10.1016/j.cmpb.2024.108561
43. Sarkar U, Fisher L, Schillinger D. Is self-efficacy associated with diabetes self-management across race/ethnicity and health literacy? *Diabetes Care.* 2006;29(4):823–829. doi:10.2337/diacare.29.04.06.dc05-1615
44. Bandura A, Freeman WH, Lightsey R. Self-efficacy: the exercise of control. *J Cognit Psychotherapy.* 1997.
45. Plaza-Zamora J, Legaz I, Osuna E, Pérez-Cárceles MD. Age and education as factors associated with medication literacy: a community pharmacy perspective. *BMC Geriatr.* 2020;20(1):501. doi:10.1186/s12877-020-01881-5
46. Tang J, Zhao Z, Guo R, et al. Preschool children's asthma medication: parental knowledge, attitudes, practices, and adherence. *Front Pharmacol.* 2024;15:1292308. doi:10.3389/fphar.2024.1292308
47. Qiao X, Tian X, Liu N, et al. The association between frailty and medication adherence among community-dwelling older adults with chronic diseases: medication beliefs acting as mediators. *Patient Educ Couns.* 2020;S0738-3991(20).
48. Turan GB, Aksoy M, Çiftçi B. Effect of social support on the treatment adherence of hypertension patients. *J Vasc Nurs.* 2019;37(1):46–51. doi:10.1016/j.jvn.2018.10.005

Patient Preference and Adherence

Dovepress
Taylor & Francis Group

Publish your work in this journal

Patient Preference and Adherence is an international, peer-reviewed, open access journal that focusing on the growing importance of patient preference and adherence throughout the therapeutic continuum. Patient satisfaction, acceptability, quality of life, compliance, persistence and their role in developing new therapeutic modalities and compounds to optimize clinical outcomes for existing disease states are major areas of interest for the journal. This journal has been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/patient-preference-and-adherence-journal>