Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Prevalence and factors associated with self-medication in Pabna, Bangladesh: An interpretable statistical and machine learning approach

Md. Shahajada Mia

shahajadabrur@gmail.com

University of Aizu

Yutaka Watanobe

University of Aizu

Protiva Rani Kundu

Pabna University of Science and Technology

Afroja Parvin

Pabna University of Science and Technology

Md. Shamim Reza

Pabna University of Science and Technology

Md. Mostafizer Rahman

University of Notre Dame

Research Article

Keywords: Self-medication, Antibiotic self-medication, Chronic condition, Multivariable logistic regression, LASSO, SHAP, Interpretable machine learning, Explainable artificial intelligence, Bangladesh

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11526128/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Prevalence and factors associated with self-medication in Pabna, Bangladesh: An interpretable statistical and machine learning approach

Md. Shahajada Mia[1,2*], Yutaka Watanobe[1], Protiva Rani Kundu[2], Afroja Parvin[2], Md. Shamim Reza[2], Md. Mostafizer Rahman[3]

[1]Graduate Department of Computer and Information Systems, The University of Aizu, Aizuwakamatsu, Fukushima 965-8580, Japan.
[2]Department of Statistics and Data Science, Pabna University of Science and Technology, Pabna-6600, Bangladesh.
[3]Lucy Family Institute for Data & Society, University of Notre Dame, IN, USA.

*Corresponding author(s). E-mail(s): shahajadabrur@gmail.com;
Contributing authors: yutaka@u-aizu.ac.jp; protivakundu3@gmail.com; afrojaparvin0000@gmail.com; shamim.reza@pust.ac.bd; mrahman3@nd.edu;

Abstract

Background: Self-medication (SM), the use of medicines without a doctor's supervision, is common in Bangladesh and other low- and middle-income countries and may cause adverse effects and contribute to antibiotic resistance. Factors associated with SM have mostly been identified with a single analytical method, leaving their consistency across methods underexplored. This study aimed to estimate the prevalence of SM and identify associated factors by comparing several methods within one interpretable framework.

Methods: A cross-sectional survey of 421 adults was conducted in Pabna Sadar, Bangladesh. Four methods were applied in parallel to 13 variables and compared at the variable level: chi-square tests, multivariable logistic regression, LASSO stability selection, and SHAP-based importance from machine learning. Nine models were compared by nested cross-validation (CV) and evaluated on a held-out test set. Factors supported by at least two methods were considered associated. Antibiotic SM was examined with multivariable logistic regression.

Results: The prevalence of SM in the previous month was 54.2%, and pharmacies were the main source of medicine information (55.7%). Among self-medicators,

39.0% took antibiotics, more often those aged 48 to 58 years than those aged 18 to 28 years. A chronic condition, gender, and education were supported by all four methods, and occupation and awareness of the antibiotic-dairy interaction by two. Respondents with a chronic condition had over four times the adjusted odds of SM (adjusted odds ratio [AOR] 4.36; 95% confidence interval [CI] 2.61 to 7.27), and males had about twice the odds of females (AOR 2.13; 95% CI 1.17 to 3.87). Random forest achieved the highest discrimination (CV AUC 0.818 ± 0.052; test AUC 0.827).

Conclusions: SM was common in this setting. The identified factors depended partly on the analytical method, and agreement across methods distinguished robust from method-dependent factors. People with chronic conditions may need particular attention, and the pharmacy counter may be a practical point for intervention, including antibiotic sales. The framework may be applied to other survey-based health studies.

Keywords: Self-medication, Antibiotic self-medication, Chronic condition, Multivariable logistic regression, LASSO, SHAP, Interpretable machine learning, Explainable artificial intelligence, Bangladesh

1 Background

Identifying the factors associated with a health behavior can help direct interventions to the groups that may benefit most, and survey data are a common source for this task. When studies of the same behavior report different factors, however, it may be unclear which groups should be prioritized. This is the case for self-medication (SM), the practice of diagnosing and treating one's own illnesses without seeking professional medical advice [1], for which the reported associated factors are often inconsistent [2, 3]. SM is practiced worldwide, and reported prevalence varies widely [4–6]: a large multi-country analysis estimated 34.3% across the European Union [3], while smaller populations reported far higher figures, such as 67% among undergraduate students in Somalia [7]. Although SM offers convenience and cost savings [8], it carries known risks, including misdiagnosis, adverse effects, drug interactions, and dependence on frequently misused medicines [9]. Antibiotic SM is also widely reported [10] and is of particular concern given its contribution to antimicrobial resistance [11].

These risks are of particular concern in low- and middle-income countries, where medicines such as antimicrobials are often obtained without a prescription [11], health-care facilities can be far from home [5], and the sale of antibiotics without a prescription may be weakly regulated [12]. Analgesics and antipyretics are among the most commonly self-medicated classes [7, 9, 13]. Bangladesh exemplifies this pattern, with SM prevalence ranging from 41% to 73.6% [2, 5, 12–14]. In the Chittagong Hill Tracts, nearly half of respondents self-medicated in the preceding year, and antibiotics were among the most commonly self-medicated drugs (34.5%) [5], while metronidazole (50.43%) was the most commonly self-medicated antibiotic in Rajshahi City [15]. SM was commonly reported for fever, cold or flu, gastric problems, pain, allergies, and headaches [13, 14]. Across studies, pharmacies, previous prescriptions, and personal

experience are frequent sources of medicines and information [5, 16], and risky practices such as not checking expiry dates have also been reported [17].

The factors associated with SM also differ across studies. For example, age, sex, and educational attainment are commonly reported socio-demographic factors [3, 14, 18], while perceived mildness of illness, prior experience with a medicine, and time savings are commonly reported reasons [2, 6, 9]. However, another study reported no significant socio-demographic differences [19]. A long-standing health condition has also been associated with SM [2], although it has received less attention than socio-demographic characteristics. Where directions can be compared, they are not always consistent: higher educational attainment has been associated with greater SM in some studies [3, 14], and lower in another [2]. Age shows a similar pattern: SM was highest among adults aged 25 to 44 in one study [3], but lower among younger adults in others [2, 14, 18]. From published results alone, it is difficult to tell whether these differences reflect real population differences or the analytical choices of each study.

Many existing studies follow a similar conventional statistical approach: bivariate analysis via the chi-square test or univariate logistic regression (LR), commonly followed by multivariable LR [5, 14, 17, 18]. A few studies have used other methods; for example, Tohan et al. [1] used one-way ANOVA and structural equation modeling to examine SM practices among university students in Bangladesh, and Nyarko et al. [20] used a maximum difference model to rank the health conditions that prompt antimicrobial SM in Ghana. LR yields interpretable adjusted odds ratios (AORs), but a single fitted model does not show whether a variable is retained under resampling, and it assumes additive effects unless interactions are specified. As a result, when studies report different factor sets, the results do not indicate which association is more robust.

Machine learning (ML) may offer part of the remedy, and it is increasingly applied in healthcare and epidemiological research [21], for example in disease prediction using ensembles of classifiers [22], including stacking ensembles that combine several base models through a meta-learner [23]. Because many ML models act as black boxes, explainable artificial intelligence methods have been developed to interpret their outputs, as reviewed by Adadi and Berrada [24] and, for medical applications, by Tjoa and Guan [25]. Among these, SHapley Additive exPlanations (SHAP), introduced by Lundberg and Lee [26], is widely used to identify influential predictors, often after several models are compared, and the best is selected [27, 28]. In Bangladesh, ML with SHAP has been applied to national survey data, for example, to identify factors for child undernutrition [29] and maternal postnatal care utilization [27]. Penalized regression has also been used for feature selection, for example, by repeatedly fitting elastic net models on data subsets to retain stable features [30].

In applied health research, statistical and ML methods are often combined in sequence, with penalized regression selecting variables, a model fitted on them, and SHAP explaining its output [28, 31], or, with LR, tree-based models, and SHAP applied one after another, as in a hybrid framework [32]. Shahid et al. [33] found that ML-enhanced and conventional LR models identified different risk factors from the same survey data. Applications of ML to SM, however, remain limited, with examples including retail purchasing data [34] and related work on medication adherence [35].

Thus, in these studies, statistical and ML methods are often applied in sequence or in isolation rather than in parallel, so their outputs are rarely compared systematically. These outputs are also not directly comparable: because categorical variables enter as dummy indicators, regression, least absolute shrinkage and selection operator (LASSO), and SHAP each return results per encoded category, while the chi-square test gives a single result per variable. The methods also have different weaknesses. Chi-square does not adjust for confounding, regression depends on a single model specification, LASSO stability selection shows how often a variable is retained but not direction or magnitude, and SHAP does not test significance. A factor identified by several of these methods may therefore be less likely to depend on any single method. In addition, antibiotic SM has more often been described than analyzed as an outcome with its own correlates (e.g., [15]).

To address these gaps, this study aims to identify the factors associated with SM using an interpretable framework that combines statistical and ML methods and compares their results at the variable level. In addition, the prevalence, patterns, and reasons for SM are estimated, and antibiotic SM is examined as a separate outcome using multivariable LR. Because the design is cross-sectional, all findings are interpreted as associations rather than as causes. The main contributions of this study are as follows.

- An interpretable framework is presented that applies bivariate chi-square testing, multivariable LR, LASSO stability selection, and SHAP-based importance in parallel to a common variable set. The framework aggregates their outputs from category to variable level so that all four are comparable, and treats agreement between at least two methods as the basis for identifying an associated factor, with agreement of all four as the strongest support, rather than relying on a single method.
- Using this framework, factors associated with SM are identified using each method separately, and the results are then compared. The variables cover demographic, socioeconomic, health, and healthcare-access and medication-related characteristics, and SHAP is derived from the best-performing of nine ML candidate models (seven classifiers and two ensembles) compared under nested cross-validation (CV). Disagreement between methods is also examined, as it may help identify associations that depend on adjustment.
- The prevalence and patterns of SM among adults in Pabna Sadar are reported, including the socio-demographic and health characteristics of respondents, the conditions self-treated, the medicines most commonly used, and the reasons reported.
- Antibiotic SM among those who self-medicate is examined as a separate outcome, and its associated factors are identified using multivariable LR.

The rest of the paper is organized as follows. Section 2 describes the study design, data collection, and the proposed framework. Section 3 presents the experimental results. Section 4 interprets the major findings, discusses their implications, and outlines the study's limitations. Finally, Section 5 concludes the paper.

2 Methods

This section presents the data collection, preprocessing, and the statistical and ML methods used to determine the prevalence and identify the associated factors of SM among adults in Pabna Sadar, Bangladesh. Fig. 1 summarizes the proposed framework.

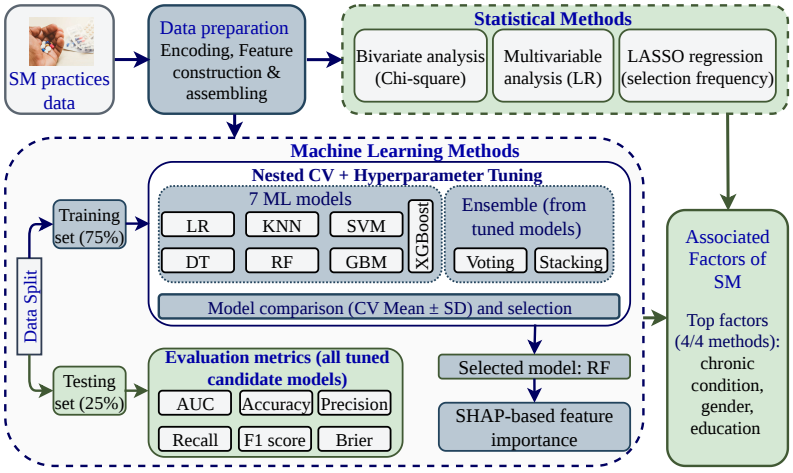


Fig. 1: Proposed interpretable framework for identifying factors associated with SM across four methods. Factors flagged by at least two methods were considered associated.

2.1 Study design and participants

This cross-sectional study was conducted among adults aged 18 years and above in Pabna Sadar, Bangladesh, from November to December 2023. Given wide cross-population variation in reported SM prevalence (34.3%–67%) [3, 7], the expected proportion was taken from a Bangladeshi community study that reported a prevalence of 60.2% [2], giving $P = 0.60$. The sample size was calculated using Cochran's formula [36], a commonly applied approach in prevalence-based health surveys [7]:

$$n = \frac{Z^2_{\alpha/2} P(1-P)}{e^2} \tag{1}$$

where $Z_{\frac{\alpha}{2}}$ is the value of the standard normal variate ($Z_{0.025} = 1.96$), and $e = 0.05$ is the acceptable margin of error. This yielded $n \approx 369$; to account for possible non-response, the sample was increased, and 421 respondents were surveyed. Convenience sampling was used, as it has been widely adopted in SM studies [7, 9, 18]. Participants who were accessible and willing were recruited via personal interviews and an online form (Google Forms).

2.2 Study variables and measurement

Data were collected using a structured questionnaire developed after reviewing the literature on SM practices. Categorical responses (e.g., yes/no items, five-point scales) were coded numerically, and SM practice was used as the response variable, coded 1 if the respondent practiced SM in the previous month and 0 otherwise. Explanatory variables were drawn from four domains: demographic (gender, residence, age group); socioeconomic (level of education, occupation, monthly family income in Bangladeshi Taka (BDT)); health characteristics (body mass index (BMI) category, presence of chronic health condition); and healthcare access and medication-related factors (perceived cost-saving of SM, doctor's advice availability, reading instructions before use, time constraint for doctor visit, awareness of antibiotic-dairy interaction). Presence of chronic health condition was a composite indicator, coded 1 if the respondent reported any of twelve listed conditions (hypertension, diabetes, gastritis, enteritis, chronic obstructive pulmonary disease, asthma, fatty liver disease, heart disease, stroke, nephritis, malignant tumour, or viral hepatitis).

Age was grouped into four 10-year intervals (18–28, 28–38, 38–48, 48–58), with the oldest respondents combined into a single open-ended category (58 and above), giving five age groups. BMI was classified using the World Health Organization (WHO) standard cutoff [37] as underweight (< 18.50), normal ($18.50 - 24.99$), overweight ($25.00 - 29.99$), or obese (≥ 30.00 kg/m^2)[1] . Two height or weight entry errors were corrected, and missing weight ($n = 8$) and height ($n = 2$) were mean-imputed. In addition, respondents' reasons for practicing SM, the diseases/symptoms self-treated, and the medicines commonly used were recorded as multi-response checklist items and summarized descriptively; these were not included as explanatory variables in the association analyses, although antibiotic use derived from the medicines item was analyzed as a secondary outcome. Five self-medicating respondents did not complete these items and were excluded from the corresponding analyses ($n = 223$).

2.3 Statistical analysis

Descriptive statistics were computed for recorded variables, including the reasons for SM, the diseases/symptoms self-treated, and the medicines most commonly used, summarized as frequencies and percentages and presented using bar diagrams to show the scenario of SM practice. The explanatory variables were then analyzed using the following statistical methods to identify factors associated with SM. Statistical significance was set at $p < 0.05$ unless otherwise stated.

Bivariate analysis. The association between each explanatory variable and SM practice was assessed using the chi-square test [38], testing the null hypothesis of independence with the test statistic:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{2}$$

[1] Overweight and obese categories were merged into a single category for analysis due to the small number of respondents in the obese category ($n = 14$).

where O_{ij} and E_{ij} are the observed and expected cell frequencies; the χ^2 statistic follows a χ^2 distribution with $(r-1)(c-1)$ degrees of freedom, where r and c are the numbers of rows and columns of the contingency table. The Cochran-Armitage test was also used to assess a linear trend in SM across the number of chronic conditions (0 to 4) [39].

Multivariable logistic regression. To assess multivariable association, multivariable LR [40] was used to obtain AORs with 95% confidence intervals (CIs), controlling for potential confounding among predictors:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \tag{3}$$

where π_i is the probability that respondent i practices SM, β_0 is the intercept, β_j are the coefficients for the explanatory variables $x_j (j = 1, 2, \cdots, k)$. Each categorical variable entered the model as a set of dummy (one-hot) indicators relative to a reference category. Crude odds ratios (CORs) were obtained from single-predictor models, and the significance of individual coefficients was assessed by Wald tests. Model fit and discrimination were then assessed using a calibration plot and the area under the receiver operating characteristic curve (AUC), and multicollinearity among predictors was evaluated using the generalized variance inflation factor (GVIF) [41]. A secondary model applied the same approach to antibiotic SM among self-medicators, defined as non-prescribed use of any antibiotic, including metronidazole, and adjusted for gender, age group, level of education, and presence of any chronic condition. This model was analyzed separately and was not part of the cross-method framework.

LASSO stability selection. To further assess which predictors were consistently retained under penalization, LASSO regression [42] was applied with stability selection [43], imposing an L_1 penalty on the regression coefficients to shrink weaker predictors toward zero:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[-\ell(\beta) + \lambda \sum_j |\beta_j| \right] \tag{4}$$

where $\hat{\beta}$ is the estimated coefficient vector, $\ell(\beta)$ is the log-likelihood of the LR model, and λ is the regularization (shrinkage) parameter controlling the penalty strength (implemented as $C = \frac{1}{\lambda} = 0.1$, with all design-matrix columns standardized before fitting). The model was refitted across 500 subsamples, each comprising 50% of the analytic sample drawn without replacement. Because each categorical predictor was represented by dummy terms, a term was counted as selected in a given subsample if its coefficient was non-zero, and a variable if any of its terms was selected; a variable was deemed stable when its variable-level selection frequency was at least 0.60.

2.4 Machine learning methods

The ML analysis followed the steps shown in Fig. 1. The dataset was first divided into 75% training and 25% testing sets using a stratified split. The same explanatory variables used in the statistical analysis were used to train seven ML models including LR classifier [40], k-nearest neighbors (KNN) [44], support vector machine (SVM) [45],

decision tree (DT) [46], random forest (RF) [47], gradient boosting machine (GBM) [48], and extreme gradient boosting (XGBoost) [49] representing diverse algorithmic families (linear, instance-based, kernel-based, tree-based, and ensemble methods). Categorical predictors were one-hot encoded (dummy coded) within a preprocessing pipeline fit only on the training data. For each ML model, hyperparameters were tuned via randomized search within a 5-fold inner CV loop (Table 1), nested inside a 5×5 repeated stratified outer CV (25 outer folds) to allow a fair comparison and reduce the optimistic bias that arises when hyperparameters are tuned and performance is estimated on the same data [50]. For each model, mean ± SD (standard deviation) across the 25 outer folds was reported for AUC, accuracy, precision, recall, F1 score, and Brier score, defined in Eqs (5)–(10); all are higher-is-better except the Brier score. AUC was used as the primary metric because it is independent of the classification threshold.

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \tag{5}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7}$$

$$F1 \text{ score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives and false negatives, respectively.

$$\text{AUC} = \frac{1}{n_+ n_-} \sum_{i=1}^{n_+} \sum_{j=1}^{n_-} \mathbf{1}\left\{\hat{\pi}(x_i^+) > \hat{\pi}(x_j^-)\right\} \tag{9}$$

where n_+ and n_- denote the number of positive and negative instances, and $\hat{\pi}(x)$ is the predicted probability for observation x; and $\mathbf{1}\{\cdot\}$ is the indicator function, equal to 1 when the enclosed condition holds and 0 otherwise. AUC is thus the probability that a randomly chosen positive instance is ranked above a randomly chosen negative one [51]. The accuracy of the predicted probabilities was assessed with the Brier score:

$$\text{Brier score} = \frac{1}{N} \sum_{i=1}^{N} (\hat{\pi}_i - y_i)^2 \tag{10}$$

where $\hat{\pi}_i$ is the predicted probability and $y_i \in \{0, 1\}$ is the observed outcome for observation i, and N is the number of observations in the evaluation set [52].

Two ensemble models were built from the seven tuned base models, each fit with its best configuration: (i) a soft-voting ensemble, weighted by a softmax of each base model's mean CV score across five metrics [53], and (ii) a stacking ensemble, in which out-of-fold predicted probabilities of the seven base models were combined by an LR meta-learner [54]. Both ensembles were evaluated under the same outer CV as the

Table 1: Hyperparameter settings and selected parameters of each ML model.

Model	Hyperparameter (search space)	Selected value
LR	C ∈ {0.001, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100}	0.03
	class_weight ∈ {None, balanced}	None
	solver ∈ {lbfgs, liblinear}	liblinear
KNN	n_neighbors ∈ {3, 5, 7, 9, 11, 15, 21}	21
	weights ∈ {uniform, distance}	distance
	p ∈ {1, 2}	1
SVM	C ∈ {0.01, 0.1, 0.3, 1, 3, 10, 30, 100}	3
	gamma ∈ {scale, auto, 0.001, 0.01, 0.1, 1}	scale
	kernel ∈ {rbf, linear}	rbf
DT	max_depth ∈ {2, 3, 4, 5, 6, 8}	8
	min_samples_leaf ∈ {3, 5, 10, 15, 20}	10
	min_samples_split ∈ {2, 5, 10, 20}	5
	criterion ∈ {gini, entropy}	entropy
RF	n_estimators ∈ {100, 200, 300, 500, 800}	300
	max_depth ∈ {3, 4, 5, 6, 8, 10, None}	None
	min_samples_leaf ∈ {1, 2, 3, 5, 8, 10}	1
	max_features ∈ {sqrt, log2, None}	log2
GBM	n_estimators ∈ {100, 200, 300, 500}	100
	learning_rate ∈ {0.01, 0.05, 0.1, 0.2}	0.1
	max_depth ∈ {2, 3, 4, 5}	5
	subsample ∈ {0.6, 0.8, 1.0}	0.8
	min_samples_leaf ∈ {1, 5, 10}	1
XGBoost	n_estimators ∈ {100, 200, 300, 500}	500
	max_depth ∈ {2, 3, 4, 5, 6}	6
	learning_rate ∈ {0.01, 0.03, 0.05, 0.1, 0.2}	0.1
	subsample ∈ {0.6, 0.8, 1.0}	0.8
	colsample_bytree ∈ {0.6, 0.8, 1.0}	0.8
	reg_lambda ∈ {0.1, 1, 5, 10}	0.1

Note: Selected values correspond to the highest cross-validated AUC per model. The number of configurations sampled by the randomized search was 25 (LR), 28 (KNN), 30 (SVM), 30 (DT), 40 (RF), 40 (GBM), and 50 (XGBoost).

seven base models (nine candidates in total), with base-model hyperparameters fixed at the values tuned on the full training set. The Friedman test [55] was then applied to assess overall differences among the nine candidates' AUC scores, followed by Wilcoxon signed-rank tests [56] with Holm correction [57] for pairwise comparisons, given the distribution-free nature of these tests. The base model with the highest mean CV AUC was selected using its configuration tuned on the full training set (Table 1). All nine candidates were evaluated once on the held-out test set, reporting the same metrics. The test AUC of the selected model was compared with those of the two ensembles using DeLong's test [58]. Finally, predictor importance was assessed for the selected ML model using SHAP [26], computed on the test set with TreeExplainer [59]. The random seed was fixed at 91 for all analyses.

2.5 Proposed interpretable framework

Because each categorical predictor was represented by multiple dummy columns, per-category estimates from SHAP, multivariable LR, and LASSO were aggregated to the variable level for comparability with the bivariate chi-square test, which evaluates each variable as a whole. For SHAP, the signed values were summed across a variable's dummy columns for each test-set observation, and variable-level importance was taken as the mean absolute value of these per-observation sums. For multivariable LR, a single overall p-value per variable was obtained via a likelihood-ratio test (LRT) comparing the full model against a reduced model omitting that variable's dummy terms:

$$G^2 = 2(\ell_{full} - \ell_{reduced}) \tag{11}$$

where ℓ_{full} and $\ell_{reduced}$ are the maximized log-likelihoods of the full and reduced models. Under H_0 that all omitted coefficients are zero, G^2 follows an asymptotic χ^2 distribution with degrees of freedom equal to the number of dummy terms removed [60]. For LASSO, the variable-level selection frequency defined in Section 2.3 was used. Associated factors of SM were then identified by combining these four variable-level signals; because this approach rests on agreement between methods rather than on any single test, a deliberately liberal screen was applied to each signal. A predictor was considered a robustly associated factor when flagged by at least two of the four methods, with agreement of all four taken as the strongest support; a variable was counted as flagged when its chi-square or likelihood-ratio p-value was at most 0.10, its selection frequency was at least 0.60, or its aggregated SHAP importance was above the median. The procedure is summarized in Algorithm 1.

3 Results

This section describes the characteristics of the respondents and the patterns and reasons for SM. It also presents the factors associated with SM across the four methods of the proposed interpretable framework and their agreement at the variable level, followed by the analysis of antibiotic SM.

3.1 Characteristics of the respondents

Of the 421 participants, 228 (54.2%) reported SM in the past month (Fig. 2a). The characteristics of respondents across four domains, with the proportion practicing SM within each category, are described below and shown in Table 2.

Demographic. Most respondents were male ($n = 324$, 77.0%), and lived in urban areas ($n = 299$, 71.0%). SM was reported by a higher proportion of male (59.6%) than female (36.1%), and by broadly similar proportions of urban (53.2%) and rural (56.6%) respondents. Nearly half of the respondents were aged 18 to 28 years ($n = 197$, 46.8%), and the number of respondents generally decreased with increasing age. SM was least common in the youngest group (47.2%), and most common among those aged 28 to 38 years (65.1%), followed by those aged 58 years and above (63.4%).

Socioeconomic. The largest educational group was graduation and above ($n = 166$, 39.4%). SM was reported by 62.0% of those with graduation and above, 58.7%

Algorithm 1 Proposed interpretable framework for identifying factors associated with SM

Input: Dataset $\mathcal{D}$, variables $\mathcal{V}$, p threshold $\alpha = 0.10$, LASSO threshold $\tau = 0.60$, subsamples $B = 500$

Output: Associated factors $\mathcal{A}$, ranked by agreement score s_v

for each $v \in \mathcal{V}$ **do**
  Compute chi-square p value $p_v^{\chi^2}$
end for
Fit multivariable LR on $\mathcal{V}$
for each $v \in \mathcal{V}$ **do**
  Compute LRT p value p_v^{LRT} (Eq (11))
end for
for $b = 1$ to B **do**
  Draw a 50% subsample of $\mathcal{D}$ without replacement
  Fit L1 penalized LR ($C = 0.1$) on standardized columns
  Record v as selected if any of its dummy terms is nonzero
end for
$f_v \leftarrow$ selection frequency of v over the B subsamples
Split $\mathcal{D}$ into stratified training (75%) and test (25%) sets
Train nine candidate models under nested CV on the training set
Select the base model with the highest mean CV AUC
Fit the selected model on the full training set with its tuned hyperparameters (Table 1)
Compute SHAP values on the test set
$I_v \leftarrow$ mean over test observations of —sum of SHAP values over the dummy columns of v—
for each $v \in \mathcal{V}$ **do**
  $s_v \leftarrow \mathbf{1}\{p_v^{\chi^2} \le \alpha\} + \mathbf{1}\{p_v^{\mathrm{LRT}} \le \alpha\} + \mathbf{1}\{f_v \ge \tau\} + \mathbf{1}\{I_v > \mathrm{median}(I)\}$
end for
$\mathcal{A} \leftarrow \{v \in \mathcal{V} : s_v \ge 2\}$, ranked by s_v
return $\mathcal{A}$

of those with no formal education, and 42.5% of those educated to secondary level. Students formed the largest occupational group ($n = 138$, 32.8%) followed by business ($n = 111$, 26.4%), service holders ($n = 97$, 23.0%) and other occupations ($n = 75$, 17.8%); the proportion practicing SM ranged from 47.1% among students to 62.9% among service holders. Just over half of the participants ($n = 221$, 52.5%) had a monthly family income of BDT 10,000 to 30,000, and SM was broadly similar across income groups.

Health. Most respondents were in the normal BMI range ($n = 257$, 61.0%), and the proportion practicing SM was higher in the overweight or obese group than in the other two groups. Chronic conditions were common: 248 respondents (58.9%) reported at least one, and SM was reported by 68.1% of them compared with 34.1% of those without.

Table 2: Socio-demographic and behavioral characteristics of respondents and their associated factors with SM using bivariate and multivariable analysis.

Variable	Category	Total, n (%)	SM		Bivariate analysis		Multivariable analysis		
			Yes, n (%)	No, n (%)	χ^2	p value	COR (95% CI)	AOR (95% CI)	p value
Gender	Female	97 (23.0)	35 (36.1)	62 (63.9)	15.652	$< 0.001^{***}$	Ref.	Ref.	–
	Male	324 (77.0)	193 (59.6)	131 (40.4)			2.61 (1.63–4.18)	2.13 (1.17–3.87)	0.014^{**}
Residence	Urban	299 (71.0)	159 (53.2)	140 (46.8)	0.274	0.601	Ref.	Ref.	–
	Rural	122 (29.0)	69 (56.6)	53 (43.4)			1.15 (0.75–1.75)	1.63 (0.95–2.80)	0.076^{*}
Age group	18–28	197 (46.8)	93 (47.2)	104 (52.8)	10.565	0.032^{**}	Ref.	Ref.	–
	28–38	86 (20.4)	56 (65.1)	30 (34.9)			2.09 (1.24–3.53)	1.66 (0.80–3.44)	0.172
	38–48	56 (13.3)	28 (50.0)	28 (50.0)			1.12 (0.62–2.03)	0.92 (0.40–2.10)	0.840
	48–58	41 (9.7)	25 (61.0)	16 (39.0)			1.75 (0.88–3.47)	1.25 (0.48–3.23)	0.651
	58+	41 (9.7)	26 (63.4)	15 (36.6)			1.94 (0.97–3.88)	1.71 (0.62–4.70)	0.298
Level of education	No formal education	46 (10.9)	27 (58.7)	19 (41.3)	11.394	0.023^{**}	Ref.	Ref.	–
	Primary	42 (10.0)	24 (57.1)	18 (42.9)			0.94 (0.40–2.19)	1.01 (0.36–2.86)	0.980
	Secondary	73 (17.3)	31 (42.5)	42 (57.5)			0.52 (0.25–1.10)	1.06 (0.40–2.81)	0.913
	Higher secondary	94 (22.3)	43 (45.7)	51 (54.3)			0.59 (0.29–1.21)	1.58 (0.56–4.46)	0.385
	Graduation and above	166 (39.4)	103 (62.0)	63 (38.0)			1.15 (0.59–2.24)	2.66 (0.98–7.20)	0.054^{*}
Occupation	Student	138 (32.8)	65 (47.1)	73 (52.9)	8.147	0.043^{**}	Ref.	Ref.	–
	Service holder	97 (23.0)	61 (62.9)	36 (37.1)			1.90 (1.12–3.23)	1.68 (0.79–3.60)	0.181
	Business	111 (26.4)	66 (59.5)	45 (40.5)			1.65 (0.99–2.73)	1.93 (0.82–4.52)	0.131
	Other	75 (17.8)	36 (48.0)	39 (52.0)			1.04 (0.59–1.82)	1.02 (0.43–2.42)	0.965
Monthly family income (BDT)	<10,000	73 (17.3)	37 (50.7)	36 (49.3)	0.999	0.801	Ref.	Ref.	–
	10,000–30,000	221 (52.5)	118 (53.4)	103 (46.6)			1.11 (0.66–1.89)	1.05 (0.56–1.95)	0.888
	30,000–60,000	74 (17.6)	43 (58.1)	31 (41.9)			1.35 (0.70–2.59)	1.19 (0.55–2.57)	0.662
	60,000+	53 (12.6)	30 (56.6)	23 (43.4)			1.27 (0.62–2.58)	1.24 (0.52–2.94)	0.624
BMI category	Underweight	53 (12.6)	27 (50.9)	26 (49.1)	5.842	0.054^{*}	1.01 (0.56–1.83)	1.47 (0.71–3.03)	0.302
	Normal	257 (61.0)	130 (50.6)	127 (49.4)			Ref.	Ref.	–
	Overweight/Obese	111 (26.4)	71 (64.0)	40 (36.0)			1.73 (1.10–2.74)	1.66 (0.95–2.89)	0.074^{*}
Presence of chronic health condition	No	173 (41.1)	59 (34.1)	114 (65.9)	46.205	$< 0.001^{***}$	Ref.	Ref.	–
	Yes	248 (58.9)	169 (68.1)	79 (31.9)			4.13 (2.74–6.24)	4.36 (2.61–7.27)	$< 0.001^{***}$
Perceived cost-saving of SM	Strongly disagree	19 (4.5)	9 (47.4)	10 (52.6)	5.452	0.244	Ref.	Ref.	–
	Disagree	69 (16.4)	45 (65.2)	24 (34.8)			2.08 (0.75–5.82)	1.48 (0.42–5.22)	0.543
	Neutral	65 (15.4)	30 (46.2)	35 (53.8)			0.95 (0.34–2.65)	0.89 (0.25–3.13)	0.851
	Agree	182 (43.2)	98 (53.8)	84 (46.2)			1.30 (0.50–3.34)	0.97 (0.30–3.11)	0.959
	Strongly agree	86 (20.4)	46 (53.5)	40 (46.5)			1.28 (0.47–3.46)	1.25 (0.37–4.28)	0.718
Doctor's advice availability	No	171 (40.6)	90 (52.6)	81 (47.4)	0.176	0.675	Ref.	Ref.	–
	Yes	250 (59.4)	138 (55.2)	112 (44.8)			1.11 (0.75–1.64)	1.50 (0.92–2.45)	$0.100^{\dagger}$
Reads instructions before use	Never	120 (28.5)	67 (55.8)	53 (44.2)	1.836	0.766	Ref.	Ref.	–
	Rarely	34 (8.1)	19 (55.9)	15 (44.1)			1.00 (0.47–2.16)	0.75 (0.29–1.95)	0.552
	Sometimes	115 (27.3)	58 (50.4)	57 (49.6)			0.80 (0.48–1.34)	1.25 (0.64–2.46)	0.508
	Often	51 (12.1)	31 (60.8)	20 (39.2)			1.23 (0.63–2.39)	1.24 (0.53–2.93)	0.622
	Always	101 (24.0)	53 (52.5)	48 (47.5)			0.87 (0.51–1.49)	0.90 (0.47–1.73)	0.759
Time constraint for doctor visit	Never	111 (26.4)	59 (53.2)	52 (46.8)	3.952	0.413	Ref.	Ref.	–
	Rarely	54 (12.8)	36 (66.7)	18 (33.3)			1.76 (0.90–3.47)	2.06 (0.90–4.72)	0.087^{*}
	Sometimes	96 (22.8)	50 (52.1)	46 (47.9)			0.96 (0.55–1.66)	0.94 (0.47–1.86)	0.857
	Often	58 (13.8)	30 (51.7)	28 (48.3)			0.94 (0.50–1.78)	1.00 (0.44–2.27)	0.996
	Always	102 (24.2)	53 (52.0)	49 (48.0)			0.95 (0.56–1.63)	1.32 (0.66–2.66)	0.434
Awareness of antibiotic-dairy interaction	Poor	37 (8.8)	21 (56.8)	16 (43.2)	6.407	0.269	Ref.	Ref.	–
	Fair	18 (4.3)	10 (55.6)	8 (44.4)			0.95 (0.31–2.96)	0.90 (0.23–3.49)	0.878
	Good	92 (21.9)	58 (63.0)	34 (37.0)			1.30 (0.60–2.82)	1.26 (0.48–3.29)	0.636
	Very good	30 (7.1)	17 (56.7)	13 (43.3)			1.00 (0.38–2.63)	1.77 (0.53–5.90)	0.355
	Excellent	17 (4.0)	6 (35.3)	11 (64.7)			0.42 (0.13–1.36)	0.21 (0.05–0.83)	0.027^{**}
	Not applicable	227 (53.9)	116 (51.1)	111 (48.9)			0.80 (0.40–1.60)	0.73 (0.31–1.73)	0.477

Note: COR = crude odds ratio; AOR = adjusted odds ratio; CI = confidence interval; Ref. = reference category. Bivariate p values are from the chi-square test; multivariable p values are Wald p values from the multivariable LR model. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.001$. $^{\dagger}p = 0.100$, borderline significance at the 10% level.

Healthcare access and medication-related. Doctor's advice was available to 250 respondents (59.4%). Most respondents agreed that SM reduces expenditure ($n = 268$, 63.7% agreeing or strongly agreeing), while 20.9% disagreed or strongly disagreed, and 15.4% were neutral. Reading instructions before taking a medicine was

inconsistent: 36.1% of respondents did so always or often, 27.3% sometimes, and 36.6% rarely or never. Time constraints on visiting a doctor were reported as always or often by 160 respondents (38.0%), sometimes by 22.8%, and rarely or never by 39.2%. Awareness of the antibiotic-dairy interaction was rated good or better by 139 respondents (33.0%) and fair or poor by 13.1%, while 227 (53.9%) chose the "not applicable" option. For the four ordinal items, the proportion practicing SM showed no consistent gradient across categories.

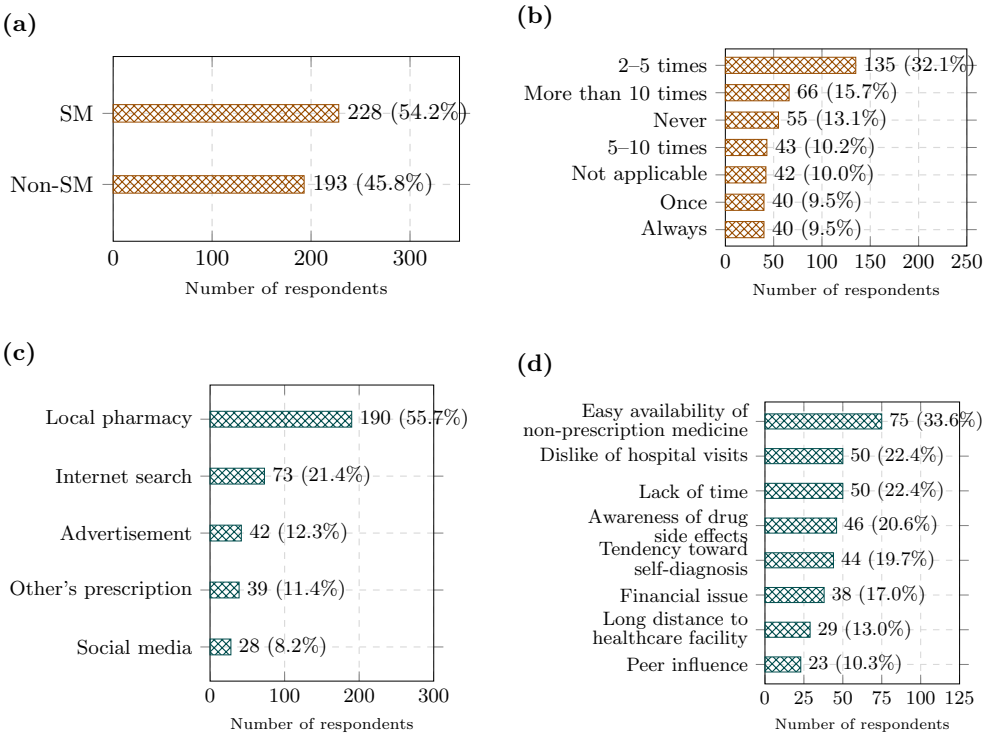


Fig. 2: Patterns of SM practice among respondents. (a) SM practice in the past month ($n = 421$). (b) Frequency of SM use in the past year ($n = 421$). (c) Sources of information for SM among past-year self-medicators ($n = 341$). (d) Reasons for SM among self-medicators ($n = 223$). Panels (c) and (d) are multiple-response items; percentages do not sum to 100%.

3.2 Patterns and reasons for SM

The frequency of SM during the previous year is shown in Fig. 2b; the most common response was two to five times ($n = 135$, 32.1%), and 42 respondents (10.0%) did not answer this item. Among the self-medicators ($n = 223$), gastric problems ($n = 104$, 46.6%) and fever, cold or flu ($n = 101$, 45.3%) were the most common

conditions treated, followed by nasal congestion (n = 71, 31.8%), and skin allergy (n = 68, 30.5%): the remaining conditions were each reported by fewer than a quarter of respondents (Supplementary information). SM was not limited to acute complaints: chronic conditions were also self-treated, including hypertension (n = 40, 17.9%) and diabetes (n = 27, 12.1%).

Asked why they treated these conditions themselves rather than seeking care, respondents often cited the availability of medicines without a prescription (n = 75, 33.6%), followed by dislike of visiting a hospital and inability to manage time (n = 50, 22.4% each); each of the remaining reasons was reported by fewer than 21% (Fig. 2d). The information behind these choices came most often from a local pharmacy (n = 190, 55.7%), followed by internet search (n = 73, 21.4%), advertisements (n = 42, 12.3%), another person's prescription (n = 39, 11.4%), and social media such as Facebook or Instagram (n = 28, 8.2%) (Fig. 2c).

Medicines reported for SM were grouped into generic classes and labeled by regulatory status using the OTC drug list proposed in Appendix 5 of the National Drug Policy 2016 [61]: classes whose medicines appear on the list were labeled OTC, those absent from it non-OTC, and those combining medicines of different status mixed (Table 3). Paracetamol (n = 146, 65.5%) was the most frequently reported OTC class. Non-OTC classes were also widely used, most often antihistamines and anti-allergics (n = 124, 55.6%), followed by guaiphenesin (n = 42, 18.8%) and montelukast (n = 28, 12.6%). Among the mixed classes, acid-reducing medicines (n = 118, 52.9%) and antibiotics (n = 87, 39.0%) were the most frequently reported; of the antibiotics, only metronidazole is listed as OTC.

Table 3: Common drug classes used for SM and their regulatory status (n = 223).

Drug class	n (%)	Status	Drug class	n (%)	Status
Paracetamol	146 (65.5)	OTC	Antidepressants	12 (5.4)	Non-OTC
Xylometazoline	25 (11.2)	OTC	Antispasmodic	7 (3.1)	Non-OTC
Multivitamin	7 (3.1)	OTC	Norethisterone	7 (3.1)	Non-OTC
Antihistamines/anti-allergics	124 (55.6)	Non-OTC	Acidity medicines[1]	118 (52.9)	Mixed
Guaiphenesin	42 (18.8)	Non-OTC	Antibiotics[2]	87 (39.0)	Mixed
Montelukast	28 (12.6)	Non-OTC	NSAIDs/analgesics[3]	61 (27.4)	Mixed
Topical antifungals/steroids	16 (7.2)	Non-OTC	Other/unclassified	5 (2.2)	–

Note: Multiple-response item; percentages do not sum to 100%.

[1]Antacids and omeprazole are listed as OTC; esomeprazole is not (22, 9.6%).

[2]Only metronidazole (15, 6.7%) is listed as OTC; all other antibiotics in this class are non-OTC.

[3]Only diclofenac gel (8, 3.6%) is listed as OTC; the other NSAIDs in this class are not.

3.3 Factors associated with SM

This subsection identifies the factors associated with SM using the proposed framework (Section 2.5). The four methods are first reported separately (Sections 3.3.1 to 3.3.4) and then compared at the variable level (Section 3.3.5).

3.3.1 Bivariate analysis

The chi-square test was used to assess the association between SM and each candidate factor (Table 2). The presence of any chronic condition ($\chi^2 = 46.21, p < 0.001$) and gender ($\chi^2 = 15.65, p < 0.001$) showed the strongest associations. Level of education ($p = 0.023$), age group ($p = 0.032$), and occupation ($p = 0.043$) were also associated at the 5% level, and BMI category was borderline, reaching the 10% significance level ($p = 0.054$), while residence and monthly family income were not associated ($p = 0.601$ and $p = 0.801$), nor were any of the five healthcare access and medication-related variables ($p \geq 0.244$). The association with chronic conditions was further examined by the number of conditions rather than presence alone, as shown in Fig. 3. SM was practiced by 59 of 173 respondents with no chronic condition (34.1%), rising to 104 of 161 with one (64.6%), 38 of 52 with two (73.1%), 20 of 26 with three (76.9%) and 6 of 7 with four (85.7%). A Cochran-Armitage test indicated a significant trend across these categories ($z = 6.42, p < 0.001$); ; two respondents reporting six chronic conditions were excluded from this test because of the sparse cell. The largest increase occurred between no condition and one, with smaller increases thereafter, although the highest categories contain few respondents, so those proportions are less precise.

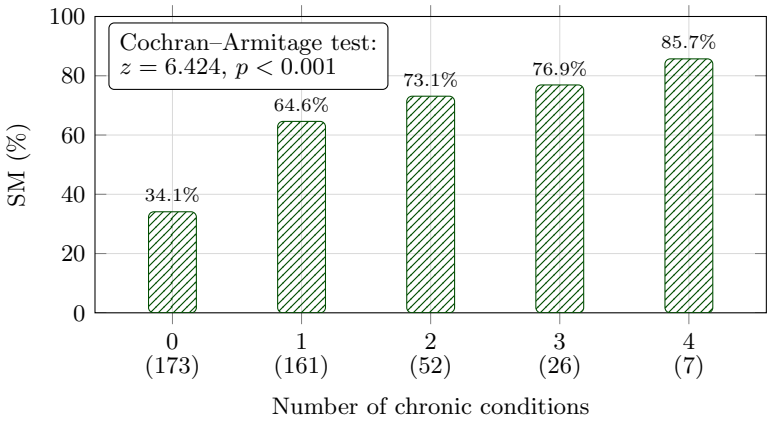


Fig. 3: Proportion of respondents practicing by number of chronic conditions. Values in parentheses on the x-axis are the number of respondents in each category.

3.3.2 Multivariable logistic regression

All thirteen explanatory variables were entered into the multivariable LR model ($n = 421,228$ events); CORs and AORs with 95% CIs and Wald p-values for every category are given in Table 2. Three category comparisons were significant at the 5% level. The odds of practicing SM were more than four times higher among respondents with at least one chronic condition than among those without (AOR 4.36, 95% CI 2.61–7.27, $p < 0.001$), the largest AOR in the model, and approximately twice as high among male as among female respondents (AOR 2.13, 95% CI 1.17–3.87, $p = 0.014$). Respondents rating their awareness of the antibiotic-dairy interaction as excellent had lower odds than those rating it poor (AOR 0.21, 95% CI 0.05–0.83, $p = 0.027$), though based on only 17 respondents.

Five further comparisons fell at or near the 10% level. The odds of practicing SM were higher among respondents with graduation and above than among those with no formal education (AOR 2.66, 95% CI 0.98–7.20, $p = 0.054$), among overweight or obese respondents than among those of normal BMI (AOR 1.66, 95% CI 0.95–2.89, $p = 0.074$), and among rural than urban residents (AOR 1.63, 95% CI 0.95–2.80, $p = 0.076$), among those who rarely rather than never reported time constraints on visiting a doctor (AOR 2.06, 95% CI 0.90–4.72, $p = 0.087$), and among those with access to doctor's advice than among those without (AOR 1.50, 95% CI 0.92–2.45, $p = 0.100$), though none of these intervals excluded unity. The time-constraint comparison was isolated, with the remaining categories of that variable sitting close to unity (AOR 1.32, 1.00, and 0.94) and no monotonic pattern across them. For residence, the adjusted estimate was larger than the crude one, suggesting an association masked in the unadjusted comparison. Within education, the intermediate categories did not differ from no formal education: primary (AOR 1.01, $p = 0.980$), secondary (AOR 1.06, $p = 0.913$), and higher secondary (AOR 1.58, $p = 0.385$).

At the variable level, the LRT identified the presence of any chronic condition, gender, and antibiotic-dairy awareness as significant at the 5% level. Level of education, residence, and availability of doctor's advice reached significance only at the 10% level ($0.05 \leq p \leq 0.10$). The remaining seven variables were not associated after adjustment (Fig. 7).

Multicollinearity was not severe: the highest $\mathrm{GVIF}^{1/(2\,\mathrm{df})}$ across the 13 predictors was 1.27 (for occupation), well below the threshold of 2.24 (Supplementary information).The model achieved an apparent AUC of 0.773, and the calibration plot (Fig. 4) showed predicted probabilities tracking observed proportions across most of the range, with the largest departures in the third and sixth deciles, indicating adequate agreement between predicted and observed probabilities.

3.3.3 LASSO stability selection

Under stability selection, four variables met the 0.60 selection-frequency threshold, namely presence of any chronic condition (1.00), gender (0.91), level of education (0.81), and awareness of antibiotic-dairy interaction (0.76). Age group fell just below the threshold at 0.56; the remaining eight variables were selected in fewer than 40%

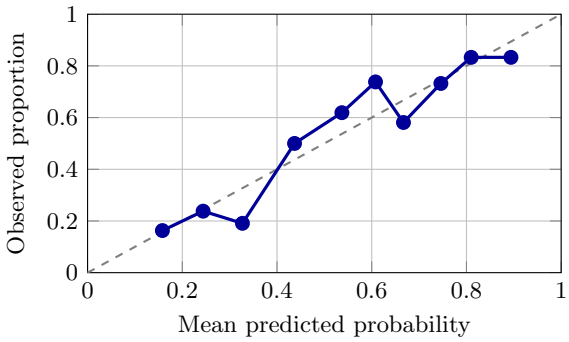


Fig. 4: Calibration plot of the multivariable LR model.

of subsamples, ranging from 0.39 for time constraints for doctor visits down to 0.11 for residence (Fig. 7; term-level frequencies in Supplementary information).

Table 4: Comparative performance of ML models under 5×5 repeated nested CV.

Model	AUC	Accuracy	Precision	Recall	F1 score	Brier score
LR	0.692 ± 0.067	0.629 ± 0.074	0.635 ± 0.071	0.796 ± 0.131	0.698 ± 0.061	0.236 ± 0.009
KNN	0.786 ± 0.055	0.706 ± 0.062	0.695 ± 0.051	**0.819** ± 0.078	0.751 ± 0.055	0.190 ± 0.022
SVM	0.774 ± 0.064	0.702 ± 0.056	0.715 ± 0.054	0.758 ± 0.075	0.734 ± 0.051	0.196 ± 0.024
DT	0.667 ± 0.071	0.641 ± 0.058	0.660 ± 0.056	0.713 ± 0.086	0.682 ± 0.053	0.250 ± 0.039
RF	**0.818** ± 0.052	**0.737** ± 0.054	**0.741** ± 0.050	0.796 ± 0.077	**0.766** ± 0.051	**0.176** ± 0.019
GBM	0.757 ± 0.053	0.697 ± 0.056	0.716 ± 0.056	0.736 ± 0.077	0.724 ± 0.054	0.235 ± 0.045
XGBoost	0.753 ± 0.070	0.683 ± 0.059	0.703 ± 0.051	0.722 ± 0.085	0.710 ± 0.059	0.222 ± 0.042
Voting	0.797 ± 0.054	0.732 ± 0.047	0.739 ± 0.044	0.785 ± 0.068	0.760 ± 0.046	0.184 ± 0.023
Stacking	0.805 ± 0.051	0.726 ± 0.057	0.732 ± 0.052	0.785 ± 0.074	0.756 ± 0.054	0.183 ± 0.021

Note: Values are reported as mean ± SD across folds; bold values indicate the best mean value for each metric (lowest for Brier score, highest for others).

3.3.4 ML model performance and SHAP importance

Seven ML classifiers (LR, KNN, SVM, DT, RF, GBM, and XGBoost) and two ensembles (soft voting and stacking) were tuned and evaluated under the nested CV scheme; their performance on six classification metrics is shown in Table 4, and the candidate and selected hyperparameters in Table 1. RF achieved the highest mean AUC across the 25 outer folds (0.818 ± 0.052), followed by the stacking ensemble (0.805±0.051) and the soft-voting ensemble (0.797±0.054), with KNN (0.786±0.055), SVM (0.774±0.064), GBM (0.757±0.053), XGBoost (0.753±0.070), LR (0.692±0.067), and DT (0.667±0.071). RF also showed the highest accuracy (0.737), precision (0.741), and F1 score (0.766), and the lowest Brier score (0.176); its recall (0.796) was second to that of KNN (0.819). The Friedman test indicated significant differences among

the nine candidates ($\chi^2 = 154.29, p < 0.001$), and raw and Holm-corrected pairwise Wilcoxon signed-rank tests indicated that RF outperformed every other candidate (all adjusted $p < 0.01$; Table 5; all 36 pairwise comparisons in Supplementary information).

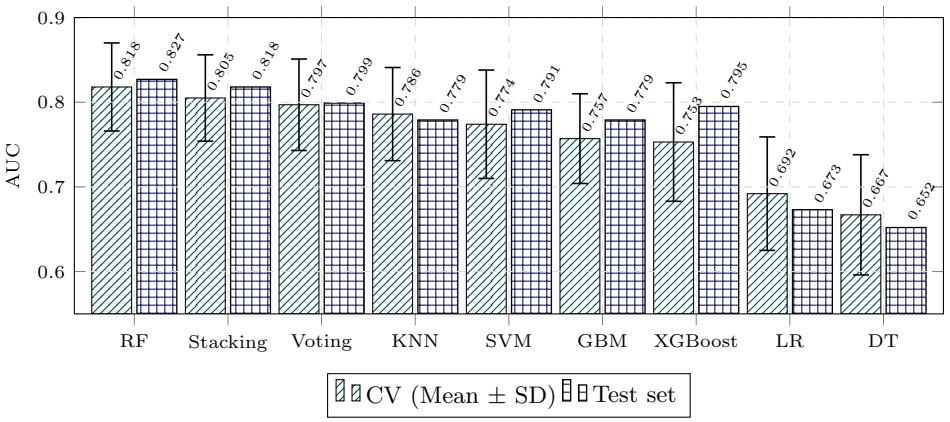


Fig. 5: Comparison of model discrimination under nested CV and held-out test evaluation; the vertical line through each CV bar indicates ± SD across the 25 outer folds.

Table 5: Pairwise comparison of model AUC against the RF using the Holm-corrected Wilcoxon signed-rank test.

RF vs. Model	ΔAUC	p (raw)	p (Holm)
LR	0.126	< 0.001	< 0.001
KNN	0.032	< 0.001	< 0.001
SVM	0.044	< 0.001	< 0.001
DT	0.151	< 0.001	< 0.001
GBM	0.060	< 0.001	< 0.001
XGBoost	0.064	< 0.001	< 0.001
Voting	0.021	< 0.001	< 0.001
Stacking	0.012	< 0.001	0.003

Note: Friedman $\chi^2 = 154.29, p < 0.001$, indicating significant differences among models. ΔAUC is the mean difference in AUC across the 25 outer folds (5×5 repeated CV); p (raw) and p (Holm) are raw and Holm-adjusted Wilcoxon signed-rank test p values. All comparisons with RF were statistically significant ($p < 0.01$).

RF was therefore selected, and all nine candidates, each fitted on the full training set with its selected hyperparameters (Table 1), were then evaluated once on the 106

respondents in the 25% test set. As shown in Fig. 5, and Table 6, RF achieved the highest AUC (0.827), with the remaining candidates ranging from 0.652 to 0.818, and the lowest Brier score (0.169); its accuracy (0.745), precision (0.788), recall (0.719), and F1 score (0.752) were below the best values, mainly obtained by SVM (0.764, 0.808, 0.737, and 0.771, respectively). However, DeLong's test showed no significant difference in test AUC between RF and the stacking ensemble ($p = 0.392$), and only a borderline difference, significant at the 10% level, in favor of RF over the soft-voting ensemble ($p = 0.058$), indicating broadly comparable discrimination on the test set; RF was therefore retained based on its CV performance. LR classifier ranked second from last (AUC 0.692 in CV and 0.673 on the test set), below the apparent AUC of the multivariable LR (0.773), as expected for out-of-sample estimates.

Table 6: Performance of ML models on the held-out test set.

Model	AUC	Accuracy	Precision	Recall	F1 score	Brier score
LR	0.673	0.670	0.690	0.702	0.696	0.231
KNN	0.779	0.689	0.700	**0.737**	0.718	0.184
SVM	0.791	**0.764**	**0.808**	**0.737**	**0.771**	0.184
DT	0.652	0.689	0.722	0.684	0.703	0.258
RF	**0.827**	0.745	0.788	0.719	0.752	**0.169**
GBM	0.779	0.708	0.771	0.649	0.705	0.207
XGBoost	0.795	0.745	0.800	0.702	0.748	0.201
Voting	0.799	0.745	0.778	**0.737**	0.757	0.182
Stacking	0.818	0.745	0.800	0.702	0.748	0.174

Note: Bold values indicate the best value for each metric (lowest for Brier score, highest for others).

SHAP values were then computed from the selected RF model on the same test set. Fig. 6a shows the twenty most influential individual levels (values for all 50 levels in Supplementary information), with each test-set respondent represented as a point and color indicating whether the level was present (red) or absent (blue). Positive SHAP values raise the predicted probability of SM and negative values lower it. Presence of at least one chronic health condition produced the largest positive SHAP values, shifting predicted probabilities upward, whereas its absence produced correspondingly negative values; gender showed the same pattern, with male respondents shifted toward higher and female respondents toward lower predicted probabilities. Among the other levels, those contributing positively included graduation and above, disagreement that SM reduces expenditure, good awareness of the antibiotic-dairy interaction, overweight or obese BMI, service holder occupation, always facing time constraints for a doctor visit, rural residence, the 28 to 38 age group, and business occupation. Those contributing negatively included student occupation, higher secondary education, the 18 to 28 age group, urban residence, and always reading instructions before use. The directions for chronic health condition, gender, residence, the highest education category, and the overweight or obese BMI category were consistent with the corresponding AORs from the multivariable LR.

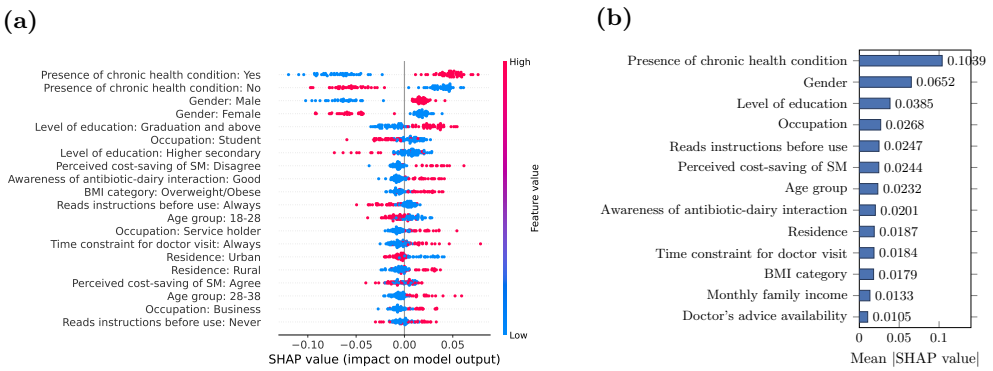


Fig. 6: SHAP-based feature importance. (a) SHAP summary (beeswarm) plot of the top 20 individual feature levels. (b) Mean absolute SHAP value by original variable.

Signed SHAP values were then aggregated across each variable's dummy columns and averaged in absolute value to obtain variable-level importance, shown in Fig. 6b (values in Supplementary information). The presence of any chronic health condition dominated the ranking (0.104), followed by gender (0.065) and level of education (0.039); occupation (0.027), reading instructions before use (0.025), and perceived cost-saving of SM (0.024) completed the six variables ranking above the median. The remaining seven ranged from 0.023 for age group down to 0.011 for availability of doctor's advice.

3.3.5 Agreement across the four methods

Combining the variable-level results of the four methods (agreement score s_v in Algorithm 1), five of the thirteen predictors were identified as factors associated with SM (Fig. 7). Agreement was complete for the presence of chronic health condition, gender, and level of education, each flagged by all four methods, and partial for occupation and awareness of antibiotic-dairy interaction, each flagged by two. The remaining eight were flagged by at most one method, and time constraint for a doctor visit and monthly family income by none. Individually, the methods flagged between four and six factors each (six by the chi-square test, six by the LRT, four by LASSO, and six by SHAP), but only three of these were common to all four methods.

The two partially supported variables were flagged by opposite pairs of methods: awareness of the antibiotic-dairy interaction was not associated in the bivariate analysis yet was significant in the multivariable LR and stable under LASSO, whereas occupation reached significance bivariately and ranked fourth by SHAP, but did not persist under either LR or LASSO, suggesting that its association may be largely accounted for by other predictors. Among the singly flagged variables, BMI category and age group were supported only by the bivariate test, residence and availability of a doctor's advice only by the LRT, and reading instructions before use and perceived cost-saving of SM only by SHAP.

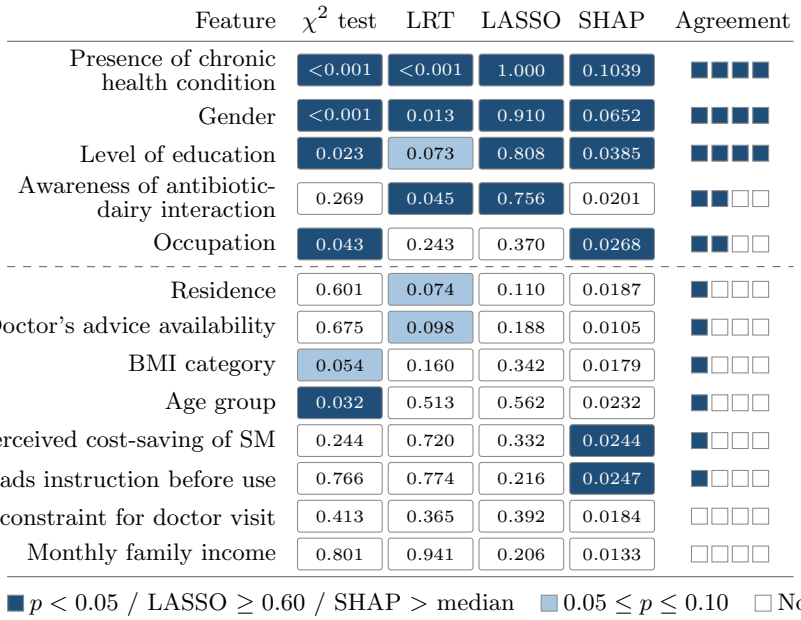


Fig. 7: Identification of associated factors with SM across the four methods. Factors are ranked by number of methods agreement (0–4); the dashed line separates factors supported by ≥ 2 methods, which were considered associated factors with SM.

3.4 Antibiotic SM and its associated factors

Antibiotics were the fourth most frequently reported drug class among self-medicators and, given the consequences of non-prescribed use for antimicrobial resistance [11], were examined separately using multivariable LR. Among the 223 self-medicators who answered the drug-use items, 87 (39.0%) used at least one antibiotic without a prescription (Table 3). Ciprofloxacin (45) and azithromycin (30) were reported most often, followed by doxycycline (19), metronidazole (15), and amoxicillin (9), while 15 respondents reported "antibiotics" without naming the agent.

Associated factors were examined within this subgroup, adjusting for gender, age group, level of education, and presence of any chronic condition; the results of which are shown in Table 7. Two category comparisons were significant at the 5% level. Among respondents who self-medicated, the odds of using an antibiotic were about three times as high for those aged 48 to 58 as for those aged 18 to 28 (AOR 3.24, 95% CI 1.13 – 9.32, $p = 0.029$). Respondents with primary education had lower odds of using antibiotics than those with no formal education (AOR 0.19, 95% CI 0.04 – 0.94, $p = 0.042$). Those with at least one chronic condition had higher odds, although this was significant only at the 10% level (AOR 1.91, 95% CI 0.94 – 3.88, $p = 0.074$). None of the remaining categories was significant: gender was not associated with antibiotic use (AOR 0.62, 95% CI 0.28 – 1.36, $p = 0.231$), and the remaining level-of-education (AOR 0.95 to 1.73) and age (AOR 0.61 to 1.57) comparisons did not approach significance.

Table 7: Factors associated with antibiotic SM among self-medicated respondents.

Variable	AOR (95% CI)	p value
Gender: Male (Ref: Female)	0.62 (0.28 – 1.36)	0.231
Age group: 28 – 38 (Ref: 18 – 28)	1.57 (0.73 – 3.39)	0.247
Age group: 38 – 48 (Ref: 18 – 28)	1.49 (0.59 – 3.75)	0.396
Age group: 48 – 58 (Ref: 18 – 28)	3.24 (1.13 – 9.32)	0.029**
Age group: 58 and above (Ref: 18 – 28)	0.61 (0.17 – 2.17)	0.448
Level of education: Primary (Ref: No formal education)	0.19 (0.04 – 0.94)	0.042**
Level of education: Secondary (Ref: No formal education)	0.95 (0.28 – 3.26)	0.939
Level of education: Higher secondary (Ref: No formal education)	1.73 (0.52 – 5.83)	0.374
Level of education: Graduation and above (Ref: No formal education)	1.66 (0.56 – 4.94)	0.365
Presence of chronic health condition: Yes (Ref: No)	1.91 (0.94 – 3.88)	0.074*

Note: $n = 223$. *$p < 0.10$; **$p < 0.05$.

4 Discussion

This section first discusses what the proposed framework revealed and how it compares with existing approaches, then interprets the associated factors, the prevalence and patterns, and antibiotic SM. It then presents the implications of these findings and outlines the limitations.

4.1 Cross-method framework findings

The main finding of the proposed framework is that the set of associated factors depended on the method used. To our knowledge, no previous study in this setting has compared four methods in this way. Used alone, each method would have identified a different set of four to six factors, and only three variables were identified by all four methods (Fig. 7), so their identification is unlikely to depend on the method used. This extends a previous observation that different LR variants identified different risk factors from the same survey data [33], and, given the inconsistent factors reported across SM studies [2, 3], suggests that the identified factors may depend partly on the analytical method.

The parallel design was central to this result. Unlike sequential pipelines [28, 31], in which a variable removed at an early step cannot be assessed by later methods, each method in this study assessed all 13 variables, so a variable rejected by one method could still be supported by others. Occupation and awareness of the antibiotic-dairy interaction, supported by opposite pairs of methods (see 3.3.5), illustrate this: a study relying on bivariate screening alone would have reported occupation and missed the antibiotic-dairy interaction; one relying on penalized selection alone would have done the opposite. The comparison therefore separated three groups of factors: those supported by all four methods; those whose association depended on adjustment, either weakening (occupation and age group) or emerging (awareness of the antibiotic-dairy interaction) after adjustment; and those supported by a single method, such as age group, which may be less robust. Two of the singly supported variables were ranked above the median by SHAP only by a small margin and without support from the

statistical methods, possibly reflecting interactions or non-linear effects that additive models do not capture, or chance variation in the small test set.

The model comparison is consistent with some non-linear structure in the data: RF (CV AUC 0.818) discriminated better than the LR classifier (0.692) with the same predictors, as did the other non-linear classifiers except the single DT (0.753 to 0.786), and neither ensemble improved on RF (Tables 4 and 5), although ensembles have been proposed to improve prediction in health datasets [22, 23]. Because RF is itself an ensemble of DTs, combining it with weaker base models may have added little in this small tabular dataset. The SHAP directions for the main factors also agreed with the corresponding AORs, indicating consistency between the ML and regression results. Reporting both agreement and disagreement between methods therefore helps show which factors are robust to the choice of method.

4.2 Interpreting the associated factors

The presence of a chronic health condition was the strongest factor by every measure. It was associated with more than four times the adjusted odds of self-medicating, was the only variable retained in all LASSO subsamples, and its SHAP importance was about 1.6 times that of gender and 2.7 times that of education (Fig. 6b). The relationship was graded rather than binary: having one chronic condition nearly doubled the proportion self-medicating (34.1% to 64.6%), while each further condition added smaller increases (Fig. 3). Chronic illness has been linked to SM [2], as has current illness [14], but the number of conditions appears to have been rarely examined in relation to SM. People who manage a long-term condition may come to know their symptoms and their medicines, so another visit may seem unnecessary.

Males had about twice the adjusted odds of self-medicating, consistent with an earlier study where women had about half the odds of men after adjustment [5], and with a crude difference of 51.2% against 31% reported elsewhere [14], although another study found no difference between men and women [2]. Rural respondents had 1.63 times the adjusted odds of urban respondents, although the interval included unity, suggesting that the association may have been masked in the crude comparison. Education did not operate as a gradient: only graduation and above was associated with higher adjusted odds, as also reported in other studies [3, 14]. This may reflect greater health literacy and greater confidence in reading one's own symptoms, which tend to come with more education, though we cannot test that here. Income was not associated, even though most respondents agreed that SM saves them money. Overweight and obese respondents self-medicated more than those of normal weight, but the difference was borderline both before and after adjustment. Time constraint and availability of a doctor's advice showed weak or no association, even though lack of time was the joint second most cited reason for self-medicating.

4.3 Prevalence and patterns compared with previous studies

The prevalence of SM in Pabna Sadar (54.2%) falls within the range reported in Bangladesh [5, 12–14]. Similar variation is reported among studies from different countries [4, 6, 16, 19]. Such variation need not reflect differences in behavior alone, because

the reference period, sampling frames, and sample sizes differ across studies, so these figures are better treated as local estimates than as national ones. The sources of and reasons for SM were also informative. Among past-year self-medicators, more than half (55.7%) obtained information about SM from local pharmacies. This is similar to other studies where participants took information from local pharmacies or pharmacists [5, 12, 19]. This reliance on local pharmacies may reflect convenience and trust in pharmacists over formal medical consultations, and it suggests that pharmacies play an important role in decision-making around SM. The reasons for SM were led by easy availability of medicines without a prescription, followed by dislike of hospital visits and time constraints (Fig. 2d), while some studies identified other reasons, including prior experience with the illness, perceiving the illness as minor, and previous familiarity with the medicine [2, 19].

Gastric acidity and cold, fever, or flu were the most common conditions treated by SM, consistent with other studies reporting fever, cold, headache, and gastric problems among the most common [4, 6, 14, 19]. Some studies provide only the broad categories of pharmaceuticals (e.g., analgesics, antibiotics) without specifying the actual drugs within them; we classified the specific medicines by regulatory status using the national OTC list [61]. Among OTC medicines, paracetamol was the most used, while antihistamines were the most used non-OTC class. Not all of this use was for minor ailments, however. Several medicines, such as the antihistamines reported, guaiphenesin, and montelukast, are not listed for sale without a prescription, and hypertension and diabetes were among the conditions people treated themselves. This may carry a risk of wrong dosing, adverse reactions, or delayed care.

4.4 Interpreting antibiotic SM

Almost two in five self-medicators reported using an antibiotic, above the 34.5% reported among self-medicators elsewhere [5]; antibiotic SM is also common globally, with a pooled prevalence of 43.0% among adults [10]. The two most named agents were ciprofloxacin and azithromycin, neither of which is listed for sale without a prescription. Such use without medical supervision is of particular concern for antimicrobial resistance [11]. More than a third of respondents rarely or never read the instructions before taking a medicine, and only a third rated their awareness of the antibiotic-dairy interaction as good or better, suggesting that medicine use was often poorly informed. Within the self-medicating group, those aged 48 to 58 had about three times the odds of using an antibiotic, and people with a chronic condition had higher odds, although this was significant only at the 10% level. Respondents with primary education had lower odds than those with no formal education (Table 7), though both estimates rest on small groups. Antibiotic use showed no significant difference by sex, although male were more likely to self-medicate overall. Taken together, these findings suggest that how antibiotics are sold needs closer attention, possibly for people managing long-term conditions.

4.5 Scalability and implications

The proposed framework is not specific to SM. It uses variables common in standard surveys, requires no additional data, and relies on standard methods available in common statistical and ML software. Its components can be adapted to other settings: another classifier can replace RF as the source of SHAP importance, and the thresholds can be adjusted to the aims of a study. It may therefore be applicable to risk-factor assessment in other survey-based studies of health behaviors, including SM, where ML has rarely been applied.

In practice, pharmacies were the main source of medicine information, and the main reason for SM was that medicines can be bought without a prescription, even though several of the medicines reported are not sold OTC. This suggests the pharmacy counter as a practical point for intervention, and it suggests that what can be sold without a prescription needs better regulation and closer oversight. Much of this dispensing is handled by drug sellers whose training in choosing medicines is limited [62], so enforcement at the point of sale may therefore matter as much as the rules themselves, and short training for those who dispense, together with a clear list of what can be sold without a prescription would help. Pharmacies and routine follow-up care may also reach those most likely to self-medicate, particularly people with a chronic condition. These findings may be relevant to other low- and middle-income countries where pharmacies are widely available and prescription rules are weakly enforced, and interventions there could focus on the characteristics most consistently associated with SM here: chronic conditions, gender, and education.

4.6 Limitations

This study has some limitations. First, the question asked whether participants had used any medication without a doctor's prescription, which may have been interpreted differently and is open to recall error. Participants may also give inaccurate answers to sensitive questions, such as those on SM status, economic conditions, and chronic conditions, and medicines were self-reported without verification. Second, as a cross-sectional study, the data were collected at one point in time, so causality cannot be established between associated factors and SM. Third, the 421 participants were from a single upazila, recruited by convenience sampling, and predominantly male, young, and educated, so the estimates may not represent the wider population. Fourth, several categories contained few respondents, so those estimates are imprecise. Finally, the framework thresholds, the single model used for SHAP, and the data split may have influenced the classification of some borderline variables, although nested CV and agreement between methods were used to reduce this dependence. Future work might use larger, multi-district samples, verify reported medicine use against pharmacy or medical records, and test the framework on data from other settings and longitudinal designs to assess generalizability.

5 Conclusions

This paper presented an interpretable framework that applies statistical and ML methods in parallel and compares their results at the variable level for identifying factors associated with SM, using survey data from adults in Pabna Sadar, Bangladesh. The presence of a chronic health condition was the strongest factor,and together with gender and education, it was supported by all four methods. Other factors were supported by only some methods, showing that relying on a single method could miss or overstate them. RF was the strongest classifier in CV, and its advantage over LR suggests that some effects may not be captured by additive models. The framework requires no additional data and may be applied to factor identification in other survey-based studies and in health informatics.

More than half of the respondents had self-medicated in the previous month, mainly because medicines were available without a prescription, and pharmacies were their main source of information. Among self-medicators, 39.0% had used antibiotics. In multivariable LR, those aged 48 to 58 years had about three times the odds of antibiotic SM compared with those aged 18 to 28 years, and people with a chronic condition had higher odds, although this was significant only at the 10% level, while antibiotic use did not differ by sex. Several of the medicines used, including most of the antibiotics reported, are not listed for sale without a prescription. Because SM can lead to side effects, delayed diagnosis, and antibiotic resistance over time, stricter rules for the sale of non-prescribed medicines, short training for those who dispense, and wider awareness among people of which medicines require a prescription may help reduce these risks.

List of abbreviations. AOR: adjusted odds ratio; AUC: area under the receiver operating characteristic curve; CI: confidence interval; COR: crude odds ratio; CV: cross-validation; GVIF: generalized variance inflation factor; LASSO: least absolute shrinkage and selection operator; LR: logistic regression; ML: machine learning; OTC: over the counter; RF: random forest; SHAP: SHapley Additive exPlanations; SM: self-medication.

Supplementary information. Additional file 1: Table S1. Diseases of self-medicated respondents ($n = 223$). Conditions self-treated by respondents who practiced SM in the previous month (multiple-response item).

Additional file 2: Table S2. GVIF for predictors in the multivariable LR model. Generalized variance inflation factors used to assess multicollinearity.

Additional file 3: Supplementary analysis results. LASSO stability selection frequencies, Friedman and Holm-corrected pairwise comparisons of model AUC, and SHAP values per level and per variable.

Declarations

Ethics approval and consent to participate. The study was approved by the Research and Technology Transfer Cell, Pabna University of Science and Technology, Bangladesh (approval number: PUST/RTTC/ECC-7). All participants provided informed consent before data collection.

Consent for publication. Not applicable.

Availability of data and materials. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests. The authors declare that they have no competing interests.

Funding. This research received no specific funding.

Authors' contributions. M.S.M.: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. Y.W.: Supervision, Validation, Writing – review & editing. P.R.K. and A.P.: Data curation, Methodology, Formal analysis. M.S.R.: Conceptualization, Visualization, Writing – review & editing. M.M.R.: Software, Validation, Writing – review & editing.

Acknowledgments. The authors thank all respondents for their time and cooperation during data collection, and the Research and Technology Transfer Cell, Pabna University of Science and Technology, Bangladesh, for its institutional support.

References

[1] Tohan, M.M., Ahmed, F., Juie, I.J., Kabir, A., Howlader, M.H., Rahman, M.A.: Knowledge, attitude, and convenience on self-medication practices among university students in Bangladesh: Exploration using structural equation modeling approach. Scientific Reports **14**(1), 10837 (2024) https://doi.org/10.1038/s41598-024-60931-9

[2] Moonajilin, M.S., Mamun, M.A., Rahman, M.E., Mahmud, M.F., Al Mamun, A.S., Rana, M.S., Gozal, D.: Prevalence and drivers of self-medication practices among savar residents in Bangladesh: A cross-sectional study. Risk Management and Healthcare Policy **13**, 743–752 (2020)

[3] Yeamans, S., Gil-de-Miguel, Á., Hernández-Barrera, V., Carrasco-Garrido, P.: Self-medication among general population in the European Union: prevalence and associated factors. European Journal of Epidemiology **39**(9), 977–990 (2024)

[4] Selvaraj, K., Kumar, S.G., Ramalingam, A.: Prevalence of self-medication practices and its associated factors in urban puducherry, India. Perspectives in clinical research **5**(1), 32–36 (2014)

[5] Saha, A., Marma, K.K.S., Rashid, A., Tarannum, N., Das, S., Chowdhury, T., Afrin, N., Chakraborty, P., Emran, M., Mehedi, H.H., *et al.*: Risk factors associated with self-medication among the indigenous communities of Chittagong hill tracts, Bangladesh. PLOS ONE **17**(6), 0269622 (2022) https://doi.org/10.1371/journal.pone.0269622

[6] Akande-Sholabi, W., Akinyemi, O.O.: Self-medication with over-the-counter drugs among consumers: a cross-sectional survey in a southwestern state

in Nigeria. BMJ Open **13**(5), 072059 (2023) https://doi.org/10.1136/bmjopen-2023-072059

[7] Moalim, S.A., Dirie, H.M., Mohamoud, J.H., Hashi, K.M., Ali, F.A., Abubakar, K.A.: Assessment of self-medication practices and its associated factors among undergraduate students of private universities in banadir region, Somalia. BMC Medical Education **26**(1), 491 (2026)

[8] World Health Organization: Guidelines for the regulatory assessment of medicinal products for use in self-medication. Technical Report WHO/EDM/QSM/00.1, World Health Organization, Geneva (2000)

[9] Cotobal-Calvo, E.M., Mata-Pérez, C., Bocchino, A., Gilart, E., Gutiérrez-Baena, B., Palazón-Fernández, J.L.: Self-medication practice and associated factors among health professionals in Spain. Nursing Reports **15**(2), 53 (2025)

[10] Gashaw, T., Yadeta, T.A., Weldegebreal, F., Demissie, L., Jambo, A., Assefa, N.: The global prevalence of antibiotic self-medication among the adult population: systematic review and meta-analysis. Systematic reviews **14**(1), 49 (2025) https://doi.org/10.1186/s13643-025-02783-6

[11] Morgan, D.J., Okeke, I.N., Laxminarayan, R., Perencevich, E.N., Weisenberg, S.: Non-prescription antimicrobial use worldwide: a systematic review. The Lancet infectious diseases **11**(9), 692–701 (2011) https://doi.org/10.1016/S1473-3099(11)70054-8

[12] Saha, A., Zam, D., Khan, A.A., Dutta, P., Mannan, A., Alam, N.: Prevalence and determinants of self-medication practices among general population: A cross-sectional study in Thimphu, Bhutan and Chattogram, Bangladesh. Journal of Public Health Research **12**(1), 22799036231152327 (2023) https://doi.org/10.1177/22799036231152327

[13] Ira, I.J.: Present condition of self-medication among general population of Comilla district, Bangladesh. The Pharma Innovation **4**(1, Part B), 87–90 (2015)

[14] Roy, N., Islam, M.N., Shahjalal, M., Siddiky, A., Imran, S.M., Aktarujjaman, M., Hossain, M.M., Rogers, B.T., Biswas, K.K., Hossain, E.: Self-medication practices among adult population in Bangladesh: A cross-sectional study. Epidemiologia **5**(2), 146–159 (2024) https://doi.org/10.3390/epidemiologia5020010

[15] Biswas, M., Roy, M.N., Manik, M.I.N., Hossain, M.S., Tapu, S.T.A., Moniruzzaman, M., Sultana, S.: Self medicated antibiotics in Bangladesh: a cross-sectional health survey conducted in the Rajshahi city. BMC Public Health **14**, 1–7 (2014) https://doi.org/10.1186/1471-2458-14-847

[16] Zeid, W., Hamed, M., Mansour, N., Diab, R.: Prevalence and associated risk factors of self-medication among patients attending el-mahsama family practice

center, Ismailia, Egypt. Bulletin of the National Research Centre **44**, 1–5 (2020) https://doi.org/10.1186/s42269-020-00351-7

[17] Tesfamariam, S., Anand, I.S., Kaleab, G., Berhane, S., Woldai, B., Habte, E., Russom, M.: Self-medication with over the counter drugs, prevalence of risky practice and its associated factors in pharmacy outlets of Asmara, Eritrea. BMC Public Health **19**, 1–9 (2019) https://doi.org/10.1186/s12889-019-6470-5

[18] Alenzi, E., Bedaiwi, S., Hamayun, R., Alanazi, A., Fawzy, M.: Key modifiable risk factors for self-medication among university students: An observational study. Exploratory Research in Clinical and Social Pharmacy **15**, 100483 (2024) https://doi.org/10.1016/j.rcsop.2024.100483

[19] Abdi, A., Faraji, A., Dehghan, F., Khatony, A.: Prevalence of self-medication practice among health sciences students in Kermanshah, Iran. BMC Pharmacology and Toxicology **19**, 1–7 (2018) https://doi.org/10.1186/s40360-018-0231-4

[20] Nyarko, E., Tchorly, F.S., Lartey, B.N.L., Chapman-Wardy, C., Ocran, E., Boiquaye, P.A.: Statistical modeling of health conditions influencing self-medication practices with antimicrobial drugs among health science students in Ghana. Discover Public Health **23**(1), 973 (2026) https://doi.org/10.1186/s12982-026-02324-z

[21] Chen, S., Yu, J., Chamouni, S., Wang, Y., Li, Y.: Integrating machine learning and artificial intelligence in life-course epidemiology: pathways to innovative public health solutions. BMC Medicine **22**(1), 354 (2024) https://doi.org/10.1186/s12916-024-03566-x

[22] Sifat, I.K., Kibria, M.K.: Optimizing hypertension prediction using ensemble learning approaches. PLoS One **19**(12), 0315865 (2024)

[23] Li, W., Peng, Y., Peng, K.: Diabetes prediction model based on ga-xgboost and stacking ensemble algorithm. PloS one **19**(9), 0311222 (2024)

[24] Adadi, A., Berrada, M.: Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access **6**, 52138–52160 (2018)

[25] Tjoa, E., Guan, C.: A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE transactions on neural networks and learning systems **32**(11), 4793–4813 (2020)

[26] Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. Advances in neural information processing systems **30** (2017)

[27] Dhar, A.K., Akther, S., Bina, F.A.: Determinants of maternal postnatal care utilization in bangladesh: A machine learning and shap-based analysis of bdhs 2022 data. Plos one **21**(5), 0350188 (2026)

[28] Meng, X., Fang, Y., Zhang, S., Huang, P., Wen, J., Peng, J., Yang, X., Ji, G., Wu, W.: Interpretable machine learning identification of dietary and metabolic factors for metabolic syndrome in southern China: a cross-sectional study. Nutrients **17**(21), 3368 (2025) https://doi.org/10.3390/nu17213368

[29] Islam, M.M., Kibria, N.M.S.J., Kumar, S., Roy, D.C., Karim, M.R.: Prediction of undernutrition and identification of its influencing predictors among under-five children in Bangladesh using explainable machine learning algorithms. PLOS ONE **19**(12), 0315393 (2024)

[30] Jenul, A., Schrunner, S., Liland, K.H., Indahl, U.G., Futsæther, C.M., Tomic, O.: RENT—repeated elastic net technique for feature selection. IEEE Access **9**, 152333–152346 (2021)

[31] Jang, J.-Y., Lee, J.-N., Park, J.-H., Lee, J.-Y.: Explainable AI-based clinical signal analysis for myocardial infarction classification and risk factor interpretation. Signals **6**(4), 62 (2025) https://doi.org/10.3390/signals6040062

[32] Khatun, M.R., Mim, M.A., Tasin, M.M., Hossain, M.M.: A hybrid framework of statistical, machine learning, and explainable AI methods for school dropout prediction. PLOS ONE **20**(9), 0331917 (2025)

[33] Shahid, M., Yahya, M.A., Song, J., Naveed, H.M., Yuksel, S., Dincer, H., Ali, M.: Machine learning vs. traditional logistic regression: predictive performance and risk factor identification for child nutritional outcome in Pakistan. BMC public health **26**(1), 667 (2025)

[34] Davies, A., Green, M.A., Singleton, A.D.: Using machine learning to investigate self-medication purchasing in England via high street retailer loyalty card data. PLOS ONE **13**(11), 0207523 (2018) https://doi.org/10.1371/journal.pone.0207523

[35] Bohlmann, A., Mostafa, J., Kumar, M.: Machine learning and medication adherence: scoping review. JMIRx Med **2**(4), 26993 (2021) https://doi.org/10.2196/26993

[36] Cochran, W.G.: Sampling Techniques vol. 3. wiley New York, ??? (1977)

[37] World Health Organization: Obesity: preventing and managing the global epidemic. World Health Organ Tech Rep Ser **894**, 1–253 (2000)

[38] McCrum-Gardner, E.: Which is the correct statistical test to use? British Journal of Oral and Maxillofacial Surgery **46**(1), 38–41 (2008) https://doi.org/10.1016/j.bjoms.2007.09.002

[39] Armitage, P.: Tests for linear trends in proportions and frequencies. Biometrics **11**(3), 375–386 (1955)

[40] Cox, D.R.: The regression analysis of binary sequences. Journal of the Royal Statistical Society Series B: Statistical Methodology **20**(2), 215–232 (1958)

[41] Fox, J., Monette, G.: Generalized collinearity diagnostics. Journal of the American Statistical Association **87**(417), 178–183 (1992)

[42] Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology **58**(1), 267–288 (1996)

[43] Meinshausen, N., Bühlmann, P.: Stability selection. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **72**(4), 417–473 (2010)

[44] Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE transactions on information theory **13**(1), 21–27 (1967)

[45] Cortes, C., Vapnik, V.: Support-vector networks. Machine learning **20**(3), 273–297 (1995)

[46] Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees, 1st edn. Chapman and Hall/CRC, ??? (1984). https://doi.org/10.1201/9781315139470

[47] Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)

[48] Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189–1232 (2001)

[49] Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)

[50] Cawley, G.C., Talbot, N.L.: On over-fitting in model selection and subsequent selection bias in performance evaluation. The Journal of Machine Learning Research **11**, 2079–2107 (2010)

[51] Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiology **143**(1), 29–36 (1982)

[52] Brier, G.W.: Verification of forecasts expressed in terms of probability. Monthly Weather Review **78**(1), 1–3 (1950)

[53] Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons, ??? (2014)

[54] Wolpert, D.H.: Stacked generalization. Neural networks **5**(2), 241–259 (1992)

[55] Friedman, M.: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the american statistical association **32**(200),

675–701 (1937)

[56] Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics bulletin **1**(6), 80–83 (1945)

[57] Holm, S.: A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics, 65–70 (1979)

[58] DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L.: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics, 837–845 (1988)

[59] Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I.: From local explanations to global understanding with explainable AI for trees. Nature machine intelligence **2**(1), 56–67 (2020)

[60] Wilks, S.S.: The large-sample distribution of the likelihood ratio for testing composite hypotheses. The annals of mathematical statistics **9**(1), 60–62 (1938)

[61] Ministry of Health and Family Welfare: National Drug Policy 2016. Technical report, Government of the People's Republic of Bangladesh, Dhaka, Bangladesh (2016). Appendix 5, List of Over-the-Counter (OTC) Drugs (Allopathic)

[62] Ahmed, S.M., Naher, N., Hossain, T., Rawal, L.B.: Exploring the status of retail private drug shops in Bangladesh and action points for developing an accredited drug shop model: a facility based cross-sectional study. Journal of pharmaceutical policy and practice **10**(1), 21 (2017)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S1Data.xlsx
- S1Table.pdf
- S2Table.pdf