JMIR Publications
The Leading eHealth Publisher

Interpretable Machine Learning for Identifying High-Risk Depression in Men Who Have Sex with Men: A Cross-Sectional Study Based on the Sexual Minority Stress Model

Xinjing Liu, Hong Chen, Jiaxiu Liu, Tianyi Zhou, Bin Lin, Xiaoni Zhong

Submitted to: JMIR Mental Health
on: August 28, 2025

Table of Contents

Interpretable Machine Learning for Identifying High-Risk Depression in Men Who Have Sex with Men: A Cross-Sectional Study Based on the Sexual Minority Stress Model

Xinjing Liu[1] MA; Hong Chen[1] MA; Jiaxiu Liu[2] MD; Tianyi Zhou[3] MA; Bin Lin[1] MD; Xiaoni Zhong[1] MD

[1] School of Public Health, Research Center for Medicine and Social Development Chongqing Medical University Chongqing CN
[2] College of Artificial Intelligence Medicine Chongqing Medical University Chongqing CN
[3] The Second People's Hospital of Yibin Yibin CN

Corresponding Author:
Xiaoni Zhong MD

School of Public Health, Research Center for Medicine and Social Development
Chongqing Medical University
No.1 Medical College Road, Yuzhong District
Chongqing
CN

Abstract

Background: Due to unique psychosocial stressors, men who have sex with men (MSM) are at a substantially higher risk of developing depressive symptoms compared with the general population. However, commonly used depression screening tools are primarily designed for the general population and fail to adequately account for MSM-specific stressors, limiting accurate risk identification in this group. In the Chinese context, the mental health of MSM is often underestimated, and there is a lack of precise assessment tools that integrate multidimensional psychosocial variables. This study aims to develop an interpretable machine learning framework to identify key risk factors for depression and enable risk stratification, thereby detecting MSM individuals at high risk for depressive symptoms and supporting timely early intervention.

Objective: This study aimed to develop machine learning models and tools to identify MSM at high risk of depression and to assess their depression risk.

Methods: Between November and December 2022, this study recruited 1,369 MSM using a non-probability sampling approach. Based on the Minority Stress Model, candidate risk factors were collected, encompassing sociodemographic characteristics, MSM-related stigma, psychological resilience, social support, and MSM sexual identity, with depressive symptoms as the primary outcome. 3 feature selection methods combined with ten machine learning algorithms were employed to construct predictive models, and model performance was evaluated to identify the optimal model. SHAP analysis was then applied to interpret the contributions of individual risk factors. To enhance clinical applicability, the risk probabilities of the optimal model were defined as risk scores (RS), which were subsequently used for risk stratification.

Results: Following feature selection, 8 potential risk factors significantly associated with depressive symptoms were identified, which were then used to construct a depression risk identification model. Among all compared algorithms, the LightGBM model demonstrated the best performance, achieving an AUC of 0.945 and exhibiting superior overall performance across multiple evaluation metrics. Key risk factors included psychological resilience, identity superiority, experiences of stigma, and internalized homophobia. Risk stratification analysis revealed that individuals in the high-risk group had a significantly higher prevalence of depressive symptoms compared with those in the low-risk group ($P < 0.001$). Visualization tools were employed to intuitively illustrate the model's classification threshold and its potential practical utility.

Conclusions: Machine learning models can effectively identify MSM individuals at high risk for depressive symptoms by integrating multidimensional features. Risk stratification based on model-derived probability scores facilitates early detection and targeted interventions, thereby supporting personalized mental health screening for this vulnerable population.

(JMIR Preprints 28/08/2025:83147)
DOI: https://doi.org/10.2196/preprints.83147

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <a href="https:/

Original Manuscript

Interpretable Machine Learning for Identifying High-Risk Depression in Men Who Have Sex with Men: A Cross-Sectional Study Based on the Sexual Minority Stress Model

Xinjing Liu[1], MA; Hong Chen[1], MA; Jiaxiu Liu, MD[2]; Tianyi Zhou[3], MA; Bing Lin[1], MD; Xiaoni Zhong[1*], MD**

1 School of Public Health, Research Center for Medicine and Social Development, Chongqing Medical University, Chongqing, China.

2 College of Artificial Intelligence Medicine, Chongqing Medical University, Chongqing, China.

3 The Second People's Hospital of Yibin, Yibin, China.

* Correspondence: Xiaoni Zhong, email: zhongxiaoni@cqmu.edu.cn

** Second-listed Correspondence: Bing Lin, email: linbing@cqmu.edu.cn

Corresponding Author:

1. Xiaoni Zhong, MM

School of Public Health

Chongqing Medical University

No.1 Medical College Road

Yuzhong District

Chongqing 400016

China

Phone: 86 13308368059

Email: zhongxiaoni@cqmu.edu.cn

2. Bin Lin, MD

School of Public Health

Chongqing Medical University

No.1 Medical College Road

Yuzhong District

Chongqing 400016

China

Phone: 86 13527545050

Email: linbing@cqmu.edu.cn

Abstract

Background Due to unique psychosocial stressors, men who have sex with men (MSM) are at a substantially higher risk of developing depressive symptoms compared with the general population. However, commonly used depression screening tools are primarily designed for the general population and fail to adequately account for MSM-specific stressors, limiting accurate risk identification in this group. In the Chinese context, the mental health of MSM is often underestimated, and there is a lack of precise assessment tools that integrate multidimensional psychosocial variables. This study aims to develop an interpretable machine learning framework to identify key risk factors for depression and enable risk stratification, thereby detecting MSM individuals at high risk for depressive symptoms and supporting timely early intervention.

Objective This study aimed to develop machine learning models and tools to identify MSM at high risk of depression and to assess their depression risk.

Methods Between November and December 2022, this study recruited 1,369 MSM using a non-probability sampling approach. Based on the Minority Stress Model, candidate risk factors were collected, encompassing sociodemographic characteristics, MSM-related stigma, psychological resilience, social support, and MSM sexual identity, with depressive symptoms as the primary outcome. 3 feature selection methods combined with ten machine learning algorithms were employed to construct predictive models, and model performance was evaluated to identify the optimal model. SHAP analysis was then applied to interpret the contributions of individual risk factors. To enhance clinical applicability, the risk probabilities of the optimal model were defined as risk scores (RS), which were subsequently used for risk stratification.

Results Following feature selection, 8 potential risk factors significantly associated with depressive symptoms were identified, which were then used to construct a depression risk identification model. Among all compared algorithms, the LightGBM model demonstrated the best performance, achieving an AUC of 0.945 and exhibiting superior overall performance across multiple evaluation metrics. Key risk factors included psychological resilience, identity superiority, experiences of stigma, and internalized homophobia. Risk stratification analysis revealed that individuals in the high-risk group had a significantly higher prevalence of depressive symptoms compared with those in the low-risk group ($P < 0.001$). Visualization tools were employed to intuitively illustrate the model's classification threshold and its potential practical utility.

Conclusion Machine learning models can effectively identify MSM individuals at high risk for depressive symptoms by integrating multidimensional features. Risk stratification based on model-derived probability scores facilitates early detection and targeted interventions, thereby supporting personalized mental health screening for this vulnerable population.

Keywords Men who have sex with men (MSM); Depression; Mental health; Sexual Minority Stress Theory; Risk prediction; Machine learning

1 Introduction

Depression is a common mental health disorder affecting individuals across all populations, primarily characterized by a persistent low mood or a loss of interest or pleasure in activities[1]. Studies have shown that men who have sex with men (MSM) exhibit a significantly higher prevalence of depression compared with the general population[2]. Globally, the prevalence of depression among MSM is approximately 35%, with the highest rates observed in Asia, reaching up to 37%[3]. In China, the prevalence of depression or depressive symptoms among this population ranges from 19.0% to 60.2%[4].

Influenced by Confucian filial piety and parental reproductive expectations, contemporary sexual minorities in China often experience conflicts between their sexual identity and societal roles[5]. This cultural tension exacerbates negative labeling and stigmatization[5], further increasing the psychological burden on MSM populations. Compared with the general population, MSM are more likely to experience adverse mental health outcomes such as depression, anxiety, and suicidal behaviors[6]. Depression, in particular, not only substantially reduces quality of life[7], but is also frequently associated with increased high-risk sexual behaviors[8], thereby elevating the risk of sexually transmitted infections, including HIV[9, 10]. Timely identification and intervention for depressive symptoms can reduce engagement in high-risk sexual behaviors and thereby contribute to controlling HIV transmission among MSM[11, 12].

The Sexual Minority Stress Model provides a key theoretical framework for understanding the mechanisms underlying depression among MSM[13]. According to the model, in addition to general environmental stressors, MSM experience stressors unique to their sexual minority status, including distal stressors (e.g., prejudice, discrimination, and violence) and proximal stressors (e.g., anticipated rejection, identity concealment, and internalized homophobia). These stressors not only directly impact mental health outcomes but are also moderated by sexual minority identity characteristics (e.g., identity salience, affective evaluation, and integration) as well as coping strategies and social support at both individual and community levels. The interplay of these factors collectively shapes the psychological health risks in MSM populations, with depression being the most common and detrimental negative outcome.

Widely used depression screening tools, such as the Center for Epidemiologic Studies Depression Scale (CES-D) [14] and the Patient Health Questionnaire-9 (PHQ-9)[15], provide effective means for assessing depression risk in the general population; however, their applicability to MSM populations is limited. These instruments were not specifically designed to capture stressors unique to MSM, such as identity conflicts[16], social stigmatization[17]and lack of social support[18], which are recognized as key contributors to depression in this population and can directly or indirectly affect mental health. Additionally, experiences of discrimination and stigma[19], historical pathologization of homosexual identity[20] and concerns about privacy and safety often discourage MSM individuals from seeking mental health services, further limiting the effectiveness of screening and intervention. Moreover, existing public health resource allocation mechanisms lack risk-based prioritization[21], hindering precise intervention for high-risk groups. Therefore, developing targeted risk assessment tools to identify and intervene with high-risk individuals is of critical importance.

However, within the unique sociocultural context of China, the mental health challenges of MSM are often underestimated, and there is a lack of risk assessment tools and intervention strategies tailored to local cultural characteristics. Previous studies on depression risk identification have primarily relied on traditional regression methods or single machine learning algorithms, which are limited in capturing complex nonlinear relationships and potential interactions among variables, and often focus on data from demographics, neuroimaging[22], or social media behaviors[23]. Currently, there is a scarcity of research that systematically integrates multidimensional psychosocial variables to identify MSM individuals at high risk for depression. To address these limitations, this study, based on the Sexual Minority Stress Model, integrates multiple psychosocial scale data relevant to

MSM and constructs an interpretable machine learning framework to identify key risk factors for depression and perform risk stratification. This approach aims to enable early and precise identification of high-risk individuals and provide scientific evidence to inform targeted and equitable mental health interventions and resource allocation.

2 Method

2.1 Study Design

This cross-sectional observational study was conducted between November and December 2022 using a structured questionnaire among MSM. The survey covered a range of demographic characteristics and validated psychosocial scales, aiming to identify subgroups of MSM at high-risk for depressive symptoms.

2.2 Participants

A total of 1,369 MSM participants were recruited using a non-probability sampling approach, including online media promotion (e.g., WeChat, QQ), peer referral, and collaboration with non-governmental organizations (NGOs). The structured questionnaire was distributed via WeChat. Participants completed the survey anonymously and were informed that they could withdraw at any time. All responses were kept strictly confidential. To ensure data quality, 278 questionnaires were excluded for the following reasons: (1) participants were under 18 years of age; (2) completion time was less than two minutes; or (3) failed logical consistency checks. Ultimately, 1,091 valid questionnaires were retained, and each participant received 15 RMB (approximately 2.18 USD) in compensation.

2.3 Measures

2.3.1 Outcomes

All participants completed the Center for Epidemiologic Studies Depression Scale (CES-D) [14], a self-report instrument for assessing depressive symptoms. The scale employs a 4-point Likert scoring system ranging from 0 (rarely or none of the time, less than 1 day) to 3 (most or all the time, 5–7 days). The total score ranges from 0 to 60, with scores above 16 indicating the presence of depressive symptoms.

2.3.2 Predictors

22 variables reflecting individual characteristics of MSM were included, covering five domains: (1) demographic characteristics; (2) MSM-related stigma; (3) psychological resilience; (4) social support; and (5) MSM sexual identity. Demographic characteristics included age, household registration, education level, employment status, marital status, and income. MSM-related stigma was measured using the multidimensional MSM stigma scale[24], which comprises four dimensions: enacted stigma, anticipated stigma, identity concealment, and internalized homophobia. Enacted stigma and identity concealment were assessed using a 4-point scale, whereas anticipated stigma and internalized homophobia were assessed using a 5-point scale, with higher scores indicating more frequent experiences of stigma. Psychological resilience was measured using the 10-item Connor-Davidson Resilience Scale (CD-RISC-10) [25]. The scale consists of 10 items, rated on a 5-point Likert scale ranging from 1 (never) to 5 (always). The total psychological resilience score was calculated as the sum of all item scores, with higher scores indicating greater resilience. Perceived social support was measured using the Multidimensional Scale of Perceived Social Support (MSPSS) [26]. The scale consists of 12 items assessing social support from three domains: family, friends, and others. Items are rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating higher levels of perceived social support. Sexual identity among MSM was assessed using the Lesbian, Gay, and Bisexual Identity Scale (LGBIS) [27], which includes the following dimensions: concealment motivation, identity uncertainty, internalized homonegativity, difficulty with identity acceptance, acceptance concerns, identity superiority, identity centrality, and identity affirmation. Items are rated on a 5-point Likert scale ranging from 1

(strongly disagree) to 5 (strongly agree).

2.4 Model Development

The dataset was randomly divided into training and validation sets in an 8:2 ratio. The training set was used for feature selection and model development, while the validation set was reserved for evaluating model performance and conducting interpretability analyses.

2.4.1 Feature Selection

We employed an integrated feature selection strategy combining 3 distinct methods: univariate analysis, the least absolute shrinkage and selection operator (LASSO), and the Boruta algorithm. Variables identified by all three methods were considered the most robust and consistent risk factors. Only those selected by all three approaches were retained. This rigorous process ultimately yielded eight variables for inclusion in the risk identification models, ensuring both statistical relevance and model parsimony.

2.4.2 Modeling

Ten machine learning algorithms were employed to develop risk identification models for depressive symptoms among MSM: decision tree (DT), random forest (RF), logistic regression (LR), ridge regression (RR), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), support vector machine (SVM), multilayer perceptron (MLP), k-nearest neighbors (KNN), and elastic net regression (ENR). For each algorithm, a randomized grid search was conducted to determine the optimal combination of hyperparameters.

2.4.3 Model Evaluation

Model performance was evaluated using sensitivity, specificity, positive predictive value (PPV), balanced accuracy (Bal. Acc), F1 score, area under the ROC curve (AUC-ROC), and area under the precision-recall curve (PR-AUC). Ten-fold cross-validation was applied to improve generalizability and reduce overfitting. Performance heatmaps were generated to compare models and identify the most suitable one for further analysis.

2.4.4 Model Interpretation

SHapley Additive exPlanations (SHAP), a game-theoretic method, was used to interpret the contribution of each risk factor to model predictions. Analyses included SHAP summary plots to assess global feature importance, SHAP beeswarm plots to visualize feature effects across samples, SHAP dependence plots to examine relationships between features and their SHAP values as well as interactions with depression outcomes, and SHAP force plots to illustrate feature contributions at the individual prediction level.

2.5 Risk Stratification

To enhance clinical applicability, risk probabilities from the optimal machine learning model were defined as risk scores (RS). Individuals were classified into high-risk (RS ≥ cutoff) and low-risk (RS < cutoff) groups, with the cutoff explicitly determined based on model performance. Boxplots illustrated the probability distributions, while bar plots with statistical tests (e.g., chi-square) compared observed depression prevalence between groups. Individual-level risk curves were also plotted to visualize the relationship between the performance-derived threshold and observed depressive outcomes.

2.6 Sensitivity Analysis

Considering that young MSM are at higher risk of depression due to psychological vulnerability and less mature sexual identity[28, 29], the optimal model was further applied to this subgroup to assess its performance and confirm the stability of the model. The results were then compared with those of the main analysis. **Figure 1** illustrates the overall analytical workflow of this study.

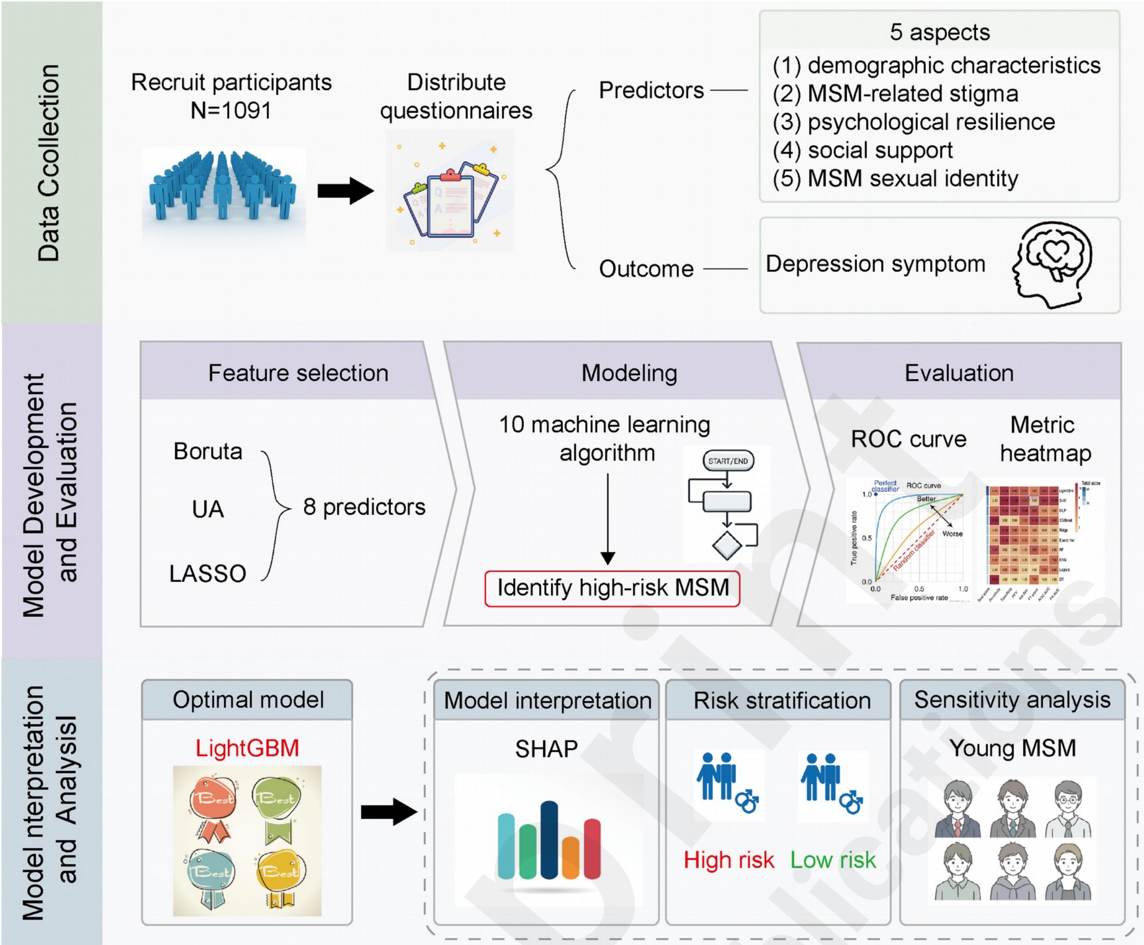


Figure1 The analysis workflow for the development and evaluation of models. UA: Univariate Analysis; LASSO: Least Absolute Shrinkage and Selection Operator. The ten machine learning algorithms employed are: decision tree (DT), random forest (RF), logistic regression (LR), ridge regression (RR), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), support vector machine (SVM), multilayer perceptron (MLP), k-nearest neighbors (KNN), and elastic net regression (ENR). Key metrics used for evaluation include: the ROC (Receiver Operating Characteristic) curve and a Metric Heatmap, which comprises multiple performance indicators such as sensitivity, specificity, positive predictive value (PPV), balanced accuracy (Bal. Acc), F1 score, area under the ROC curve (AUC-ROC), and area under the precision-recall curve (PR-AUC).

2.7 Statistical Analysis

All analyses were conducted using R software (version 4.4.3). The *glmnet*, *caret*, and *MASS* packages were applied for feature selection to implement the LASSO and Boruta algorithms. The *tidymodels* framework (a collection of R packages for modeling and machine learning) was employed to develop and evaluate various risk identification models. Data visualization was performed using the *ggplot2* package to generate clear and customizable plots. Model interpretability was assessed using the *fastshap* package for computing SHAP values, and feature-level contributions were further visualized using *ggbeeswarm*.

3 Results

3.1 Participant Characteristics

A total of 1,091 MSM participants were included in the final analysis. The demographic characteristics and depression status of the study participants are shown in **Table S1** in Multimedia Appendix. Among them, 744 participants (68.2%) were identified as having depressive symptoms (CES-D score≥ 16), while 347 (31.8%) were classified as non-depressed. **Table 1** presents the demographic characteristics and psychosocial scale scores of the study population, stratified by depression status.

Table 1 Baseline characteristics of the whole dataset

Characteristic	Total(n=1091)	Depression（n=744）	Non-depression(n =347)	P-value
Demographic factors				
Age, mean (SD)	26.34 (6.32)	26.11 (6.43)	26.84 (6.08)	.07
Household registration, n (%)				.03
City	757 (69.4)	501 (67.3)	256 (73.8)	
Town	334 (30.6)	243 (32.7)	91 (26.2)	
Education, n (%)				<.001
Illiterate and semi-illiterate	8 (0.7)	8 (1.1)	0 (0.0)	
Primary school	16 (1.5)	16 (2.2)	0 (0.0)	
Junior high school	60 (5.5)	53 (7.1)	7 (2.0)	
Senior high school/ vocational high school/ secondary technical school	179 (16.4)	145 (19.5)	34 (9.8)	
Junior college	291 (26.7)	207 (27.8)	84 (24.2)	
Bachelor's degree and above	537 (49.2)	315 (42.3)	222 (64.0)	
Employment, n (%)				<.001
Employed	728 (66.7)	480 (64.5)	248 (71.5)	
Retired	28 (2.6)	28 (3.8)	0 (0.0)	
Student	269 (24.7)	184 (24.7)	85 (24.5)	
Unemployed or jobless	66 (6.0)	52 (7.0)	14 (4.0)	
Marital status, n (%)				.01
Unmarried	761 (69.8)	502 (67.5)	259 (74.6)	
Married	293 (26.9)	210 (28.2)	83 (23.9)	
Married/ Widowed	37 (3.4)	32 (4.3)	5 (1.4)	
Income, n (%)				.04

<1000	49 (4.5)	39 (5.2)	10 (2.9)	
1000-3000	234 (21.4)	159 (21.4)	75 (21.6)	
3001-5000	344 (31.5)	246 (33.1)	98 (28.2)	
5001-10000	346 (31.7)	231 (31.0)	115 (33.1)	
>10000	118 (10.8)	69 (9.3)	49 (14.1)	
Scale scores				
Enacted stigma, mean (SD)	10.64 (4.21)	11.95 (4.05)	7.82 (3.00)	<.001
Perceived stigma, mean (SD)	16.17 (5.54)	17.27 (4.79)	13.80 (6.26)	<.001
Identity concealment, mean (SD)	9.35 (3.43)	10.27 (3.15)	7.40 (3.19)	<.001
Internalized homophobia, mean (SD)	11.37 (4.41)	12.88 (3.92)	8.12 (3.56)	<.001
Psychological resilience, mean (SD)	36.38 (8.55)	34.89 (7.80)	39.55 (9.20)	<.001
Family support, mean (SD)	14.76 (3.79)	14.15 (3.55)	16.07 (3.94)	<.001
Friend support, mean (SD)	14.98 (3.71)	14.30 (3.50)	16.44 (3.72)	<.001
Others support, mean (SD)	15.10 (3.76)	14.33 (3.54)	16.77 (3.69)	<.001
Concealment motivation, mean (SD)	10.80 (3.07)	11.08 (2.56)	10.18 (3.88)	<.001
Identity uncertainty, mean (SD)	11.79 (4.38)	13.16 (3.69)	8.84 (4.29)	<.001
Internalized homonegativity, mean (SD)	9.32 (3.22)	10.12 (2.83)	7.61 (3.33)	<.001
Difficulty with identity acceptance, mean (SD)	9.01 (2.23)	10.26 (2.46)	8.33 (2.90)	<.001
Acceptance attention, mean (SD)	10.26 (2.92)	10.73 (2.50)	9.27 (3.48)	<.001
Identity superiority, mean (SD)	7.97 (3.45)	9.10 (3.21)	5.55 (2.58)	<.001
Identity centrality, mean (SD)	13.00 (3.61)	13.64 (3.32)	11.63 (3.81)	<.001
Identity affirmation, mean (SD)	6.70 (2.17)	6.76 (2.03)	6.56 (2.45)	.15

3.2 Feature Selection

A total of 22 candidate variables were evaluated using three feature selection methods: Boruta, LASSO and UA. The selection results for each method are illustrated in **Figure 2**. Variables that were selected by at least three methods were considered as final predictors, including age, enacted stigma, Internalized homophobia, psychological resilience, others support, identity uncertainty, internalized homonegativity and identity superiority. Ultimately, these eight variables were incorporated into model development to ensure both statistical reliability and practical interpretability.

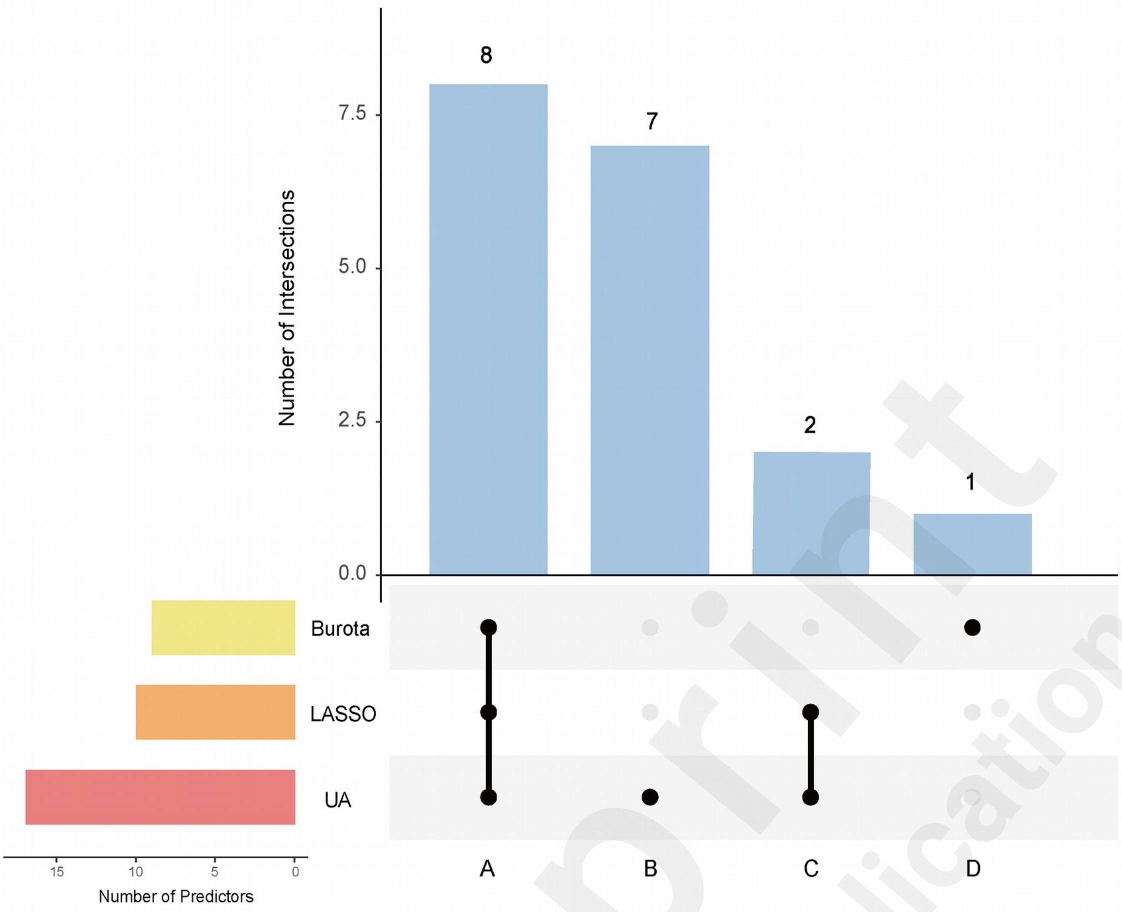


Figure 2 Upset plot of interactions between the predictors: **A** Age, Enacted stigma, Internalized homophobia, Psychological resilience, Others support, Identity uncertainty, Internalized homonegativity, Identity superiority. **B** Perceived stigma, Identity concealment, Friend support, Concealment motivation, Acceptance attention, Identity centrality, Identity affirmation. **C** Family support, Difficulty with identity acceptance **D** Household registration.

3.3 Model Performance

A comparison of demographic characteristics between the training and testing sets is presented in **Table S2** in Multimedia Appendix. As shown in **Figure 3A**, all models demonstrated strong discriminative ability, with AUC-ROC values exceeding 0.87. The support vector machine (SVM) achieved the highest AUC (0.948), followed closely by LightGBM (AUC = 0.945). **Figure 3B** presents a heatmap of performance metrics for each model, highlighting that LightGBM achieved the highest overall score, indicating the most balanced performance across all evaluation criteria. Cross-validation results in **Figure 3C** further confirm the robustness of model discrimination, with both LightGBM and SVM maintaining high AUC values. Regarding calibration, **Figure 3D** shows that

LightGBM attained the lowest Brier score (0.064), reflecting high accuracy in its probability estimates. The confusion matrices for all depression identification models are provided in **Figure S1** in Multimedia Appendix. Considering that the primary goal was to accurately identify high-risk individuals and provide reliable risk estimates, LightGBM was selected as the optimal model, reflecting a favorable balance between classification performance and calibration, and supporting its practical application in identifying MSM at elevated risk of depression.

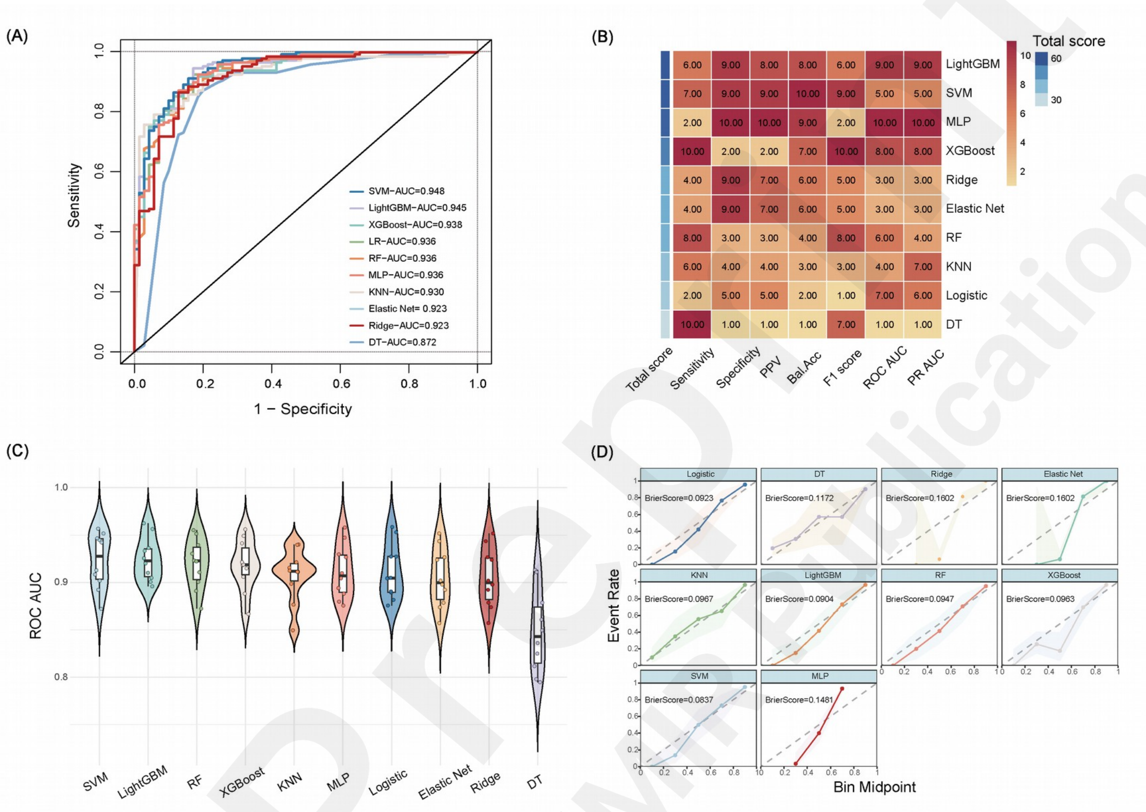


Figure 3 Performance comparison of different machine learning models for depression risk prediction: (A) Receiver operating characteristic (ROC) curves of ten machine learning models, with corresponding AUC values. (B) Heatmap ranking the performance of different models across multiple evaluation metrics, including sensitivity, specificity, PPV, balanced accuracy, F1 score, ROC AUC, and PR AUC. (C) Distribution of ROC AUC values from cross-validation for each model shown as violin plots. (D) Calibration curves for the ten models, with Brier scores indicating calibration performance.

3.4 Model Interpretation with SHAP

SHAP values were used to interpret the optimal model. Figure 3A presents global feature importance ranked by mean absolute SHAP values. Psychological resilience and identity superiority emerged as the most influential predictors, followed by enacted stigma, internalized homophobia, and others support. In contrast, features such as internalized homonegativity and age contributed relatively little. Figure 3B visualizes the direction and magnitude of SHAP values across individuals. Higher psychological resilience was associated with negative SHAP values, indicating a protective effect against depression, whereas higher identity superiority and enacted stigma acted as risk factors. Figure 3C shows SHAP dependence plots for the top four features. A nonlinear protective effect of psychological resilience was observed, particularly for scores above 40. Furthermore, identity superiority and internalized homophobia were positively associated with depression risk. Enacted stigma also exhibited a strong monotonic effect, with SHAP values increasing as the level of stigma rose. Figure 3D illustrates a SHAP force plot for a high-risk individual with a risk probability of 0.972. The main drivers of increased risk included higher identity superiority, greater internalized homophobia, and lower psychological resilience, while features such as age and internalized homonegativity contributed minimally.

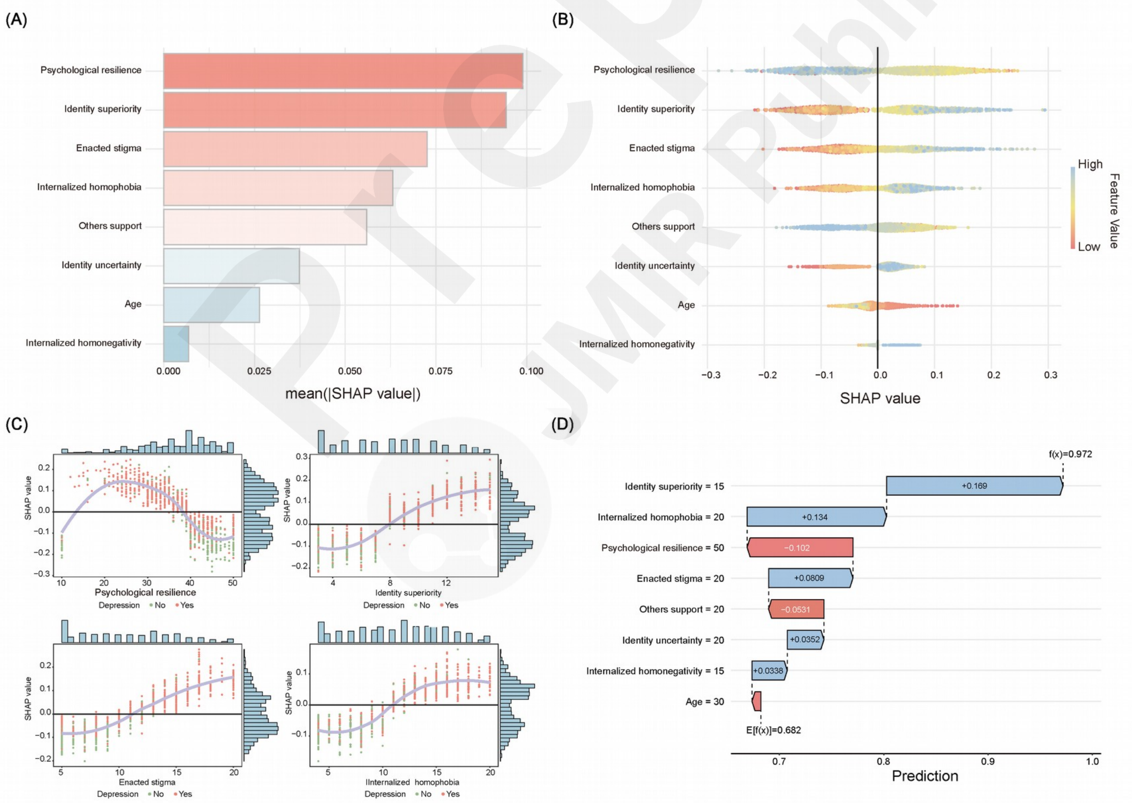


Figure 4 SHAP-based interpretation of the optimal machine learning model (LightGBM) for depression risk identification. (A) Global feature importance ranked by mean SHAP values. (B) SHAP beeswarm plot illustrating feature effects across all samples. (C) SHAP dependence plots

showing the marginal effects of the top four features (psychological resilience, identity superiority, enacted stigma, and internalized homophobia) on the prediction outcome. (D) SHAP force plot displaying feature contributions for the 125th sample.

3.5 Risk Stratification

Figure 5A depicts the distribution of risk probabilities for the high- and low-risk groups, showing a clear separation in risk scores between the two groups. **Figure 5B** compares depression screening outcomes, with 91.5% (118/129) of individuals in the high-risk group screened positive for depression, compared to only 30.0% (27/90) in the low-risk group; this difference was statistically significant ($P < 0.0001$). **Figure 5C** further illustrates risk stratification based on risk scores (RS). Using a cutoff of RS = 0.695, the high-risk group closely corresponded with actual depression screening outcomes, as visually demonstrated in the bar plot above.

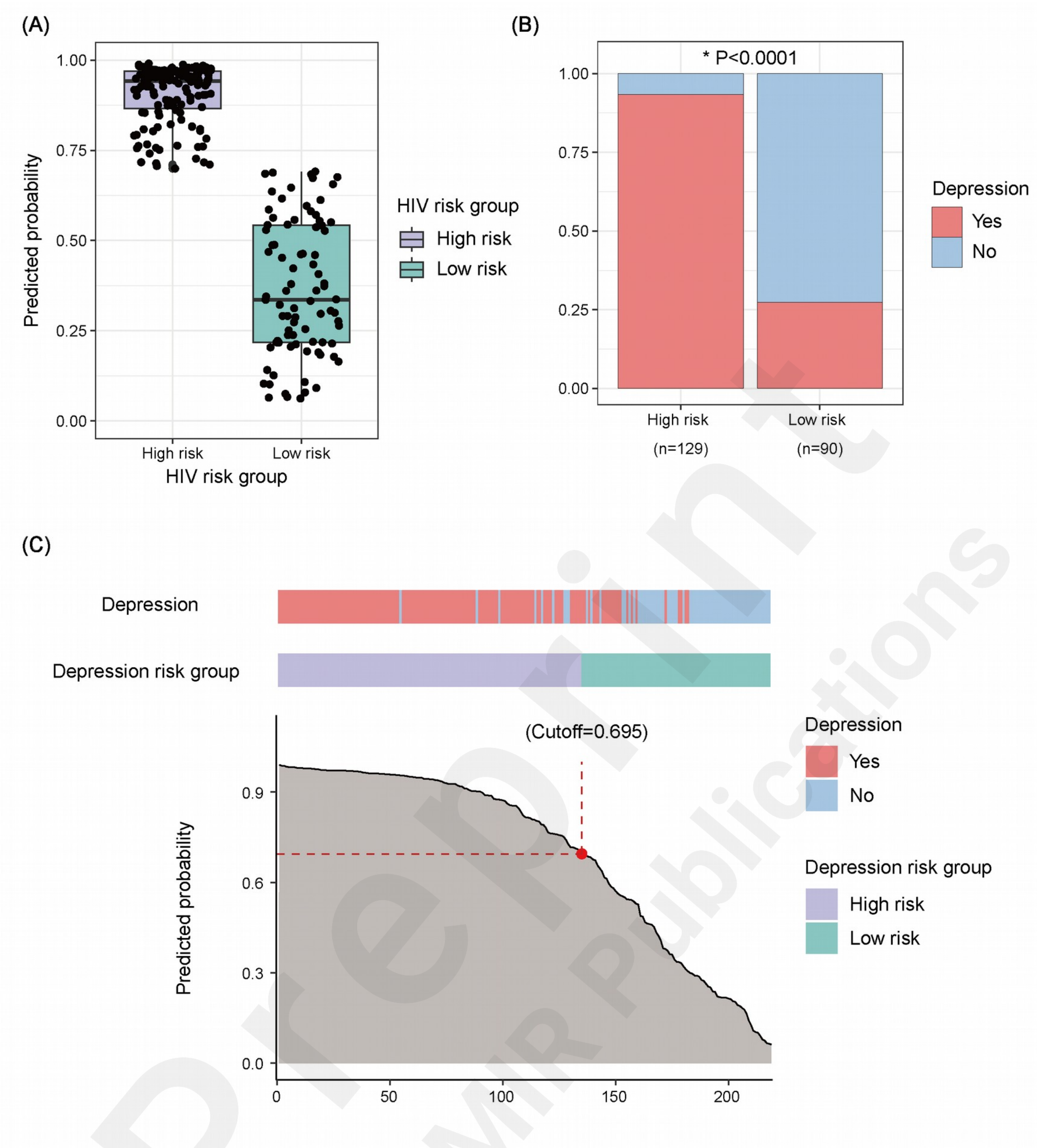


Figure 5 Evaluation of the predicted risk scores of the optimal model in the testing set. (A) Distribution of predicted probabilities and depression risk groups. (B) Comparison of actual depression prevalence between the risk groups. (C) Risk stratification plot showing the optimal cutoff (0.695) for separating high- and low-risk groups, with individual depression outcomes displayed above.

3.6 Sensitivity analysis

The optimal model was further applied to the young MSM dataset, yielding an AUC of 0.943, which indicates that the model was also effective in identifying high-risk individuals within this subgroup (**Figure S2** in Multimedia Appendix). Moreover, comparisons between young MSM and the training and testing sets (**Tables S3–S4** in Multimedia Appendix) showed no significant differences, except for age.

4 Discussion

4.1 principal findings

Based on the minority stress model[13]□this study comprehensively considered both proximal and distal stressors, integrating multidimensional variables such as internalized homophobia, social support, and experiences of stigma. 10 machine learning algorithms were employed to develop classification models aimed at identifying individuals at high risk of depressive symptoms. The results indicated that the models effectively distinguished individuals at risk of depressive symptoms, and the derived risk scores enabled meaningful population-level risk stratification. This approach demonstrates potential as a screening tool in public health practice Notably, within the MSM population—where the prevalence of depressive symptoms is considerably higher than in the general population—the model can identify potential cases that may be overlooked by conventional methods, thereby facilitating early detection, the implementation of individualized interventions, and the development of targeted mental health strategies.

In this study, 68.2% of MSM self-reported depressive symptoms, consistent with findings from recent research[30]. In contrast, a meta-analysis of MSM studies conducted between 2004 and 2018 reported a pooled prevalence of 40.0% (95% CI: 37.9%–45.0%) [4], indicating that our observed prevalence is substantially higher and suggesting a potential increase in the psychological health burden among MSM in recent years. This discrepancy may be closely related to both temporal and sociocultural factors. On one hand, with increasing societal attention to sexual minorities, individuals may be more willing to report psychological distress accurately in surveys. On the other hand, MSM in southwestern China continue to face challenges such as uneven economic development, limited mental health resources, social discrimination, and internalized stigma. Above all, these findings further highlight the heightened vulnerability of MSM in terms of mental health and underscore the need for early identification and individualized interventions within this population.

In this study, the LightGBM model demonstrated the best performance, achieving an AUC of 0.945 and a well-balanced between performance and stability across multiple metrics. Its advantages include the ability to process high-dimensional data, capture nonlinear relationships among variables, and maintain robustness in the presence of outliers and missing values. Key predictors included psychological resilience, identity superiority, experiences of stigma, and internalized homophobia, consistent with previous research on depression risk among MSM populations[31-34].

Among the key predictors, psychological resilience emerged as the most influential variable. Previous studies have generally identified resilience as a protective factor against depression[31, 35]，but rarely quantified the impact of different resilience levels on depression risk. Our findings revealed that moderate-to-low levels of resilience provided significant protection, whereas extremely high levels might yield diminishing returns or even mask underlying psychological distress, thereby delaying help-seeking behaviors. This suggests that future interventions should prioritize fostering adaptive coping skills rather than simply pursuing ever-higher resilience. The impact of identity superiority on depression risk followed an S-shaped trajectory, in which moderate-to-high levels of self-identity were associated with a significantly increased risk of depression. This phenomenon may reflect the cognitive and sociocultural conflicts that MSM individuals encounter during the process of self-identity affirmation in the context of pervasive stigma and discrimination. At these levels of identity, individuals may become more sensitive to external negative evaluations and social exclusion, which can heighten psychological stress and anxiety, thereby increasing the likelihood of depression. In addition, experiences of stigma exhibited a positive linear association with depression risk: the more frequent or severe the exposure, the greater the psychological burden and the higher the risk of depression. Prolonged exposure to discrimination and stigmatization may undermine psychological resilience, foster negative psychological states such as internalized homophobia and self-denial, and thereby exacerbate depressive symptoms. Based on these findings, future interventions could focus on supporting identity development and alleviating stigma-related stress. Targeted approaches, such as cognitive-behavioral interventions, psychological health education, and the establishment of social support networks, may help MSM populations effectively cope with societal pressures and identity-related conflicts, ultimately reducing the risk of depression.

This study developed a risk identification model intended as an auxiliary tool to assist mental health professionals in detecting individuals at potential high risk, rather than serving as an independent diagnostic instrument. Compared with conventional one-dimensional questionnaire screening, the model demonstrated improved accuracy in identifying high-risk populations and optimizing the allocation of limited mental health resources, thereby supporting personalized intervention. Through SHAP-based interpretability analysis, the study quantified both the direction and magnitude of each feature's contribution to depression risk, shedding light on the underlying mechanisms driving model predictions and informing the prioritization and design of intervention strategies. In practice, the risk identification results should be regarded as indicative rather than deterministic and followed by professional evaluation conducted by trained specialists; clinical diagnosis should continue to rely on standardized psychiatric assessments and expert judgment. Furthermore, future research should incorporate fairness audits to evaluate predictive performance across different population subgroups, ensuring that the tool can be applied equitably and safely in real-world contexts.

4.2 Limitations

Despite the progress made in this study, several limitations should be acknowledged. First, the cross-sectional design precludes establishing temporal sequences or causal relationships among the risk factors. Second, all variables were self-reported, which may introduce information bias and compromise data accuracy. Third, both the training and validation sets were drawn from the same study sample, limiting the assessment of the model's generalizability using independent external datasets. Fourth, the study sample may not fully represent the broader MSM population, particularly individuals with limited internet access, lower educational attainment, or those residing in rural areas. This limitation may restrict the applicability of the findings to more marginalized subgroups, who may exhibit different depression risk profiles or data patterns. Future research should employ multi-center, large-scale, longitudinal studies to further evaluate the stability and generalizability of the model.

5 Conclusion

This study demonstrates that machine learning models can effectively identify MSM individuals at high risk for depressive symptoms. Among the models, LightGBM exhibited the best performance, and SHAP-based interpretability analysis provided key insights into the nonlinear relationships of critical risk factors. The developed model can be applied for risk stratification within MSM populations, offering clinicians and individuals a tool for effective risk assessment and enabling more timely intervention and prevention.

Acknowledgments

The authors would like to thank all participants in this study.

Ethics approval

This study was approved by the Ethics Committee of Chongqing Medical University (Approval No. 2019001).

Funding

This study was funded by: the National Natural Science Foundation of China (No. 72504044), the Chongqing Science and Technology Commission (Grant No. CSTB2025NSCQ-GPX1200) and the National Key Project for Infectious Diseases of the Ministry of Science and Technology of China (No. 2018ZX10721102–005).

Declaration of Competing Interest

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.

Data availability

The data that supports the findings of this study are available from the corresponding author, upon reasonable request.

Authors' Contributions

XZ and BL were responsible for the conception and design of the study. XL, HC and TZ were in charge of the analysis and interpretation and the data collection process. All authors were involved in the drafting of the manuscript and critically reviewed it for important intellectual content. All authors gave their approval for the final version to be published and agreed to be accountable for all aspects of the work.

Abbreviations

MSM	Men have sex with men
LASSO	Least Absolute Shrinkage and Selection Operator

Multimedia Appendix

The details of the a forementioned appendices are provided in the Multimedia Appendix.

Reference

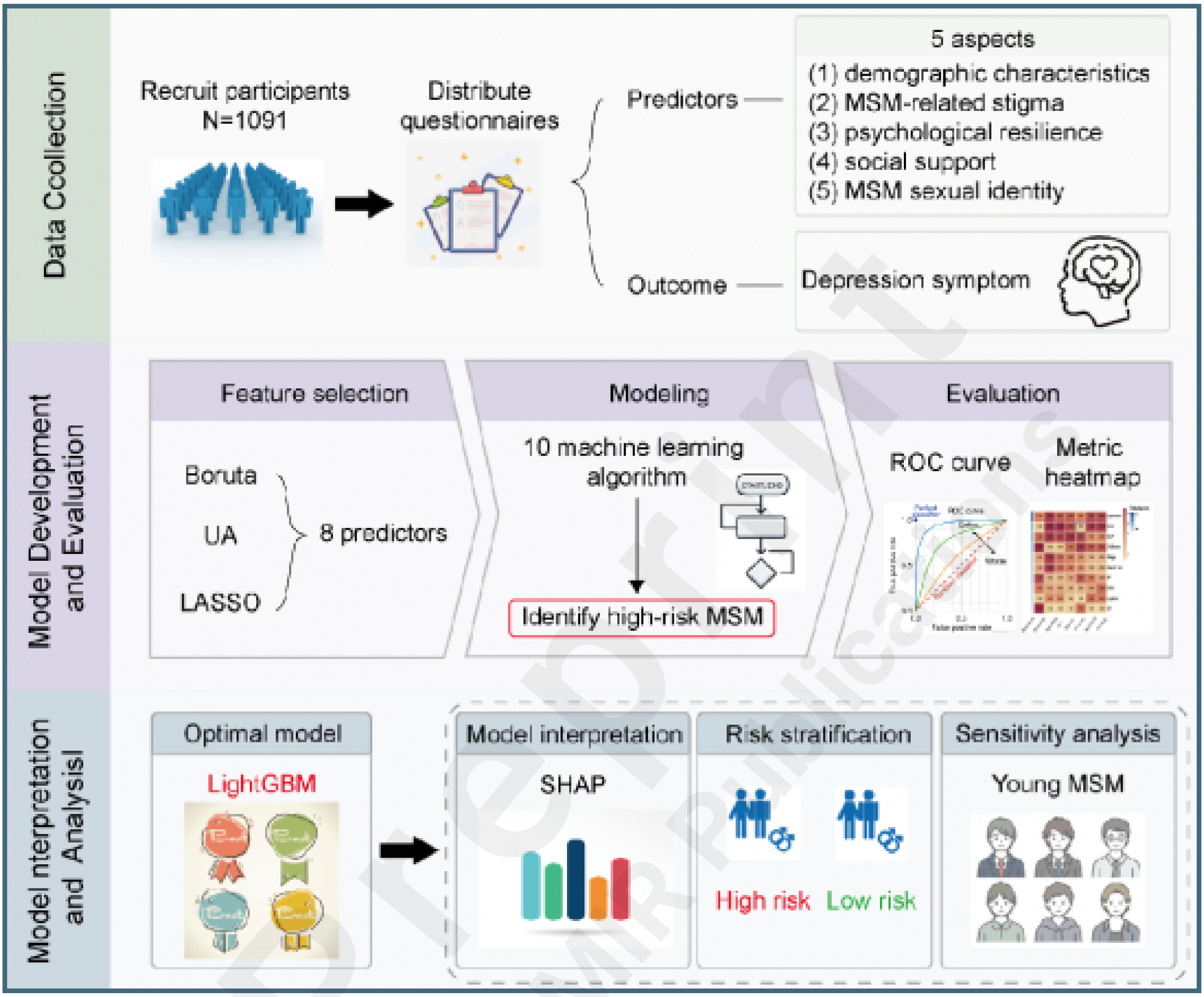
1. WHO. Depression. [updated 2025.8.3]; Available from: https://www.who.int/health-topics/depression#tab=tab_1.
2. Mills TC, Paul J, Stall R, Pollack L, Canchola J, Chang YJ, et al. Distress and Depression in Men Who Have Sex With Men: The Urban Men's Health Study. American Journal of Psychiatry. 2004 2004/02/01;161(2):278-85. doi: 10.1176/appi.ajp.161.2.278.
3. Nouri E, Moradi Y, Moradi G. What is the global prevalence of depression among men who have sex with men? A systematic review and meta-analysis. Annals of General Psychiatry. 2022 2022/09/12;21(1):38. doi: 10.1186/s12991-022-00414-1.
4. Fu H, Feng T, Wang T, Wu X, Cai Y, Yang T. Reported prevalence of depression or depressive symptoms among men who have sex with men in China, 2004–2018: A systematic review and meta-analysis. Journal of Affective Disorders. 2020 2020/12/01/;277:192-203. doi: https://doi.org/10.1016/j.jad.2020.08.011.
5. Whyke TWJJoH. Discourses of heteronormativity and power: The ethical position of Confucianism on same-sex behaviour in China. 2023;70(9):1787-806.
6. Wei D, Wang X, You X, Luo X, Hao C, Gu J, et al. Prevalence of depression, anxiety and suicide among men who have sex with men in China: a systematic review and meta-analysis. Epidemiology and Psychiatric Sciences. 2020;29:e136. doi: 10.1017/S2045796020000487.
7. Hohls JK, König HH, Quirke E, Hajek A. Anxiety, Depression and Quality of Life-A Systematic Review of Evidence from Longitudinal Observational Studies. International journal of environmental research and public health. 2021 Nov 16;18(22). PMID: 34831779. doi: 10.3390/ijerph182212022.
8. Shiu CS, Chen YC, Tseng PC, Chung AC, Wu MT, Hsu ST, et al. Curvilinear Relationship Between Depression and Unprotected Sexual Behaviors among Men Who Have Sex with Men. The Journal of Sexual Medicine. 2014;11(10):2466-73. doi: 10.1111/jsm.12638 %J The Journal of Sexual Medicine.
9. Ahaneku H, Ross MW, Nyoni JE, Selwyn B, Troisi C, Mbwambo J, et al. Depression and HIV risk among men who have sex with men in Tanzania. AIDS care. 2016;28 Suppl 1(sup1):140-7. PMID: 27002772. doi: 10.1080/09540121.2016.1146207.
10. Fendrich M, Avci O, Johnson TP, Mackesy-Amiti ME. Depression, substance use and HIV risk in a probability sample of men who have sex with men. Addictive behaviors. 2013 2013/03/01/;38(3):1715-8. doi: https://doi.org/10.1016/j.addbeh.2012.09.005.
11. O'Cleirigh C, Newcomb ME, Mayer KH, Skeer M, Traeger L, Safren SA. Moderate levels of depression predict sexual transmission risk in HIV-infected MSM: a longitudinal analysis of data from six sites involved in a "prevention for positives" study. AIDS and behavior. 2013 Jun;17(5):1764-9. PMID: 23605154. doi: 10.1007/s10461-013-0462-8.
12. Preston DB, D'Augelli AR, Kassab CD, Cain RE, Schulze FW, Starks MT. The Influence of Stigma on the Sexual Risk Behavior of Rural Men Who Have Sex with Men. AIDS Education and Prevention. 2004 2004/08/01;16(4):291-303. doi: 10.1521/aeap.16.4.291.40401.
13. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychological bulletin. 2003 Sep;129(5):674-97. PMID: 12956539. doi: 10.1037/0033-2909.129.5.674.
14. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Applied Psychological Measurement. 1977 1977/06/01;1(3):385-401. doi: 10.1177/014662167700100306.
15. Williams N. PHQ-9. Occupational Medicine. 2014;64(2):139-40. doi: 10.1093/occmed/kqt154.
16. Williams DY, Hall WJ, Dawes HC, Srivastava A, Radtke SR, Ramon M, et al. Relationships between internalized stigma and depression and suicide risk among queer youth in the United States: a systematic review and meta-analysis. Frontiers in psychiatry. 2023;14:1205581. PMID: 37547195. doi: 10.3389/fpsyt.2023.1205581.
17. Okonkwo N, Rwema JOT, Lyons C, Liestman B, Nyombayire J, Olawore O, et al. The Relationship Between Sexual Behavior Stigma and Depression Among Men Who have Sex with Men and Transgender Women in Kigali, Rwanda: a Cross-sectional Study. International journal of mental health and addiction. 2022 Dec;20(6):3228-43. PMID: 36532817. doi: 10.1007/s11469-021-00699-5.
18. Mohamad Fisal ZA, Minhat HS, Mohd Zulkefli NA, Ahmad N. Biopsychosocial approach to understanding determinants of depression among men who have sex with men living with HIV: A

systematic review. PloS one. 2022;17(3):e0264636. PMID: 35286312. doi: 10.1371/journal.pone.0264636.
19. Altman D, Aggleton P, Williams M, Kong T, Reddy V, Harrad D, et al. Men who have sex with men: stigma and discrimination. The Lancet. 2012;380(9839):439-45. doi: 10.1016/S0140-6736(12)60920-9.
20. Sun S, Pachankis JE, Li X, Operario D. Addressing Minority Stress and Mental Health among Men Who Have Sex with Men (MSM) in China. Current HIV/AIDS Reports. 2020 2020/02/01;17(1):35-62. doi: 10.1007/s11904-019-00479-w.
21. Ikhile IA, Mavhandu-Mudzusi AH. Catalyzing change: A framework for equitable health among MSM—insights from South Africa for developing nations. Sexuality, Gender & Policy. 2024 2024/05/01;7(2):122-51. doi: https://doi.org/10.1002/sgp2.12091.
22. Gao S, Calhoun VD, Sui J. Machine learning in major depression: From classification to treatment outcome prediction. 2018;24(11):1037-52. doi: https://doi.org/10.1111/cns.13048.
23. Li Y, Cai M, Qin S, Lu X. Depressive Emotion Detection and Behavior Analysis of Men Who Have Sex With Men via Social Media. Frontiers in psychiatry. 2020;11:830. PMID: 32922323. doi: 10.3389/fpsyt.2020.00830.
24. Brown CA, Sullivan PS, Stephenson R, Baral SD, Bekker L-G, Phaswana-Mafuya NR, et al. Developing and validating the Multidimensional Sexual Identity Stigma Scale among men who have sex with men in South Africa. Stigma and Health. 2021;6(3):277-86. doi: 10.1037/sah0000294.
25. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the connor–davidson resilience scale (CD-RISC): Validation of a 10-item measure of resilience. Journal of Traumatic Stress. 2007 2007/12/01;20(6):1019-28. doi: https://doi.org/10.1002/jts.20271.
26. Tonsing K, Zimet GD, Tse S. Assessing social support among South Asians: the multidimensional scale of perceived social support. Asian journal of psychiatry. 2012 Jun;5(2):164-8. PMID: 22813661. doi: 10.1016/j.ajp.2012.02.012.
27. Mohr JJ, Kendra MSJJocp. Revision and extension of a multidimensional measure of sexual minority identity: the Lesbian, Gay, and Bisexual Identity Scale. 2011;58(2):234.
28. Steever J, Francis J, Gordon LP, Lee J. Sexual minority youth. Primary care. 2014 Sep;41(3):651-69. PMID: 25124211. doi: 10.1016/j.pop.2014.05.012.
29. Poštuvan V, Podlogar T, Zadravec Šedivy N, De Leo D. Suicidal behaviour among sexual-minority youth: a review of the role of acceptance and support. The Lancet Child & adolescent health. 2019 Mar;3(3):190-8. PMID: 30679139. doi: 10.1016/s2352-4642(18)30400-0.
30. Fu J, Chen X, Dai Z, Huang Y, Xiao W, Wang H, et al. HIV-related stigma, depression and suicidal ideation among HIV-positive MSM in China: a moderated mediation model. BMC public health. 2023 Oct 27;23(1):2117. PMID: 37891525. doi: 10.1186/s12889-023-17047-y.
31. Yan H, Li X, Li J, Wang W, Yang Y, Yao X, et al. Association between perceived HIV stigma, social support, resilience, self-esteem, and depressive symptoms among HIV-positive men who have sex with men (MSM) in Nanjing, China. 2019;31(9):1069-76.
32. Li X, Yan H, Wang W, Yang H, Li SJAoE. Association between enacted stigma, internalized stigma, resilience, and depressive symptoms among young men who have sex with men in China: a moderated mediation model analysis. 2021;56:1-8.
33. Zhou T, Chen Q, Zhong XJAoE. Associations between sexual identity stigma and positive sexual identity, and depression among men who have sex with men in China: A mediation analysis using structural equation modelling. 2024;96:32-9.
34. Bruce D, Harper GW, Bauermeister JAJPoso, diversity g. Minority stress, positive identity development, and depressive symptoms: Implications for resilience among sexual minority male youth. 2015;2(3):287.
35. Chakrapani V, Vijin PP, Logie CH, Newman PA, Shunmugam M, Sivasubramanian M, et al. Understanding How Sexual and Gender Minority Stigmas Influence Depression Among Trans Women and Men Who Have Sex with Men in India. LGBT health. 2017 Jun;4(3):217-26. PMID: 28422615. doi: 10.1089/lgbt.2016.0082.

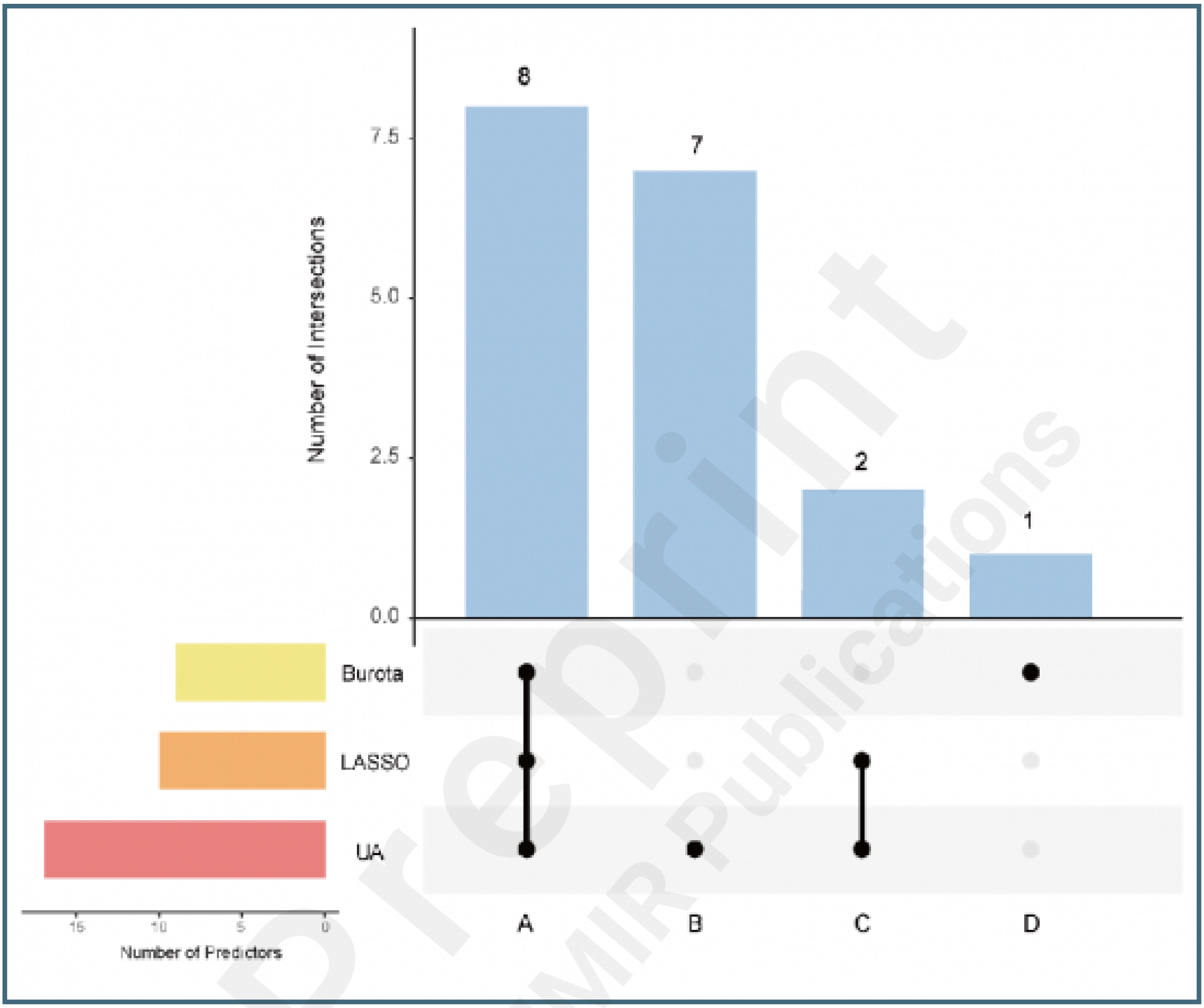
Supplementary Files

Figures

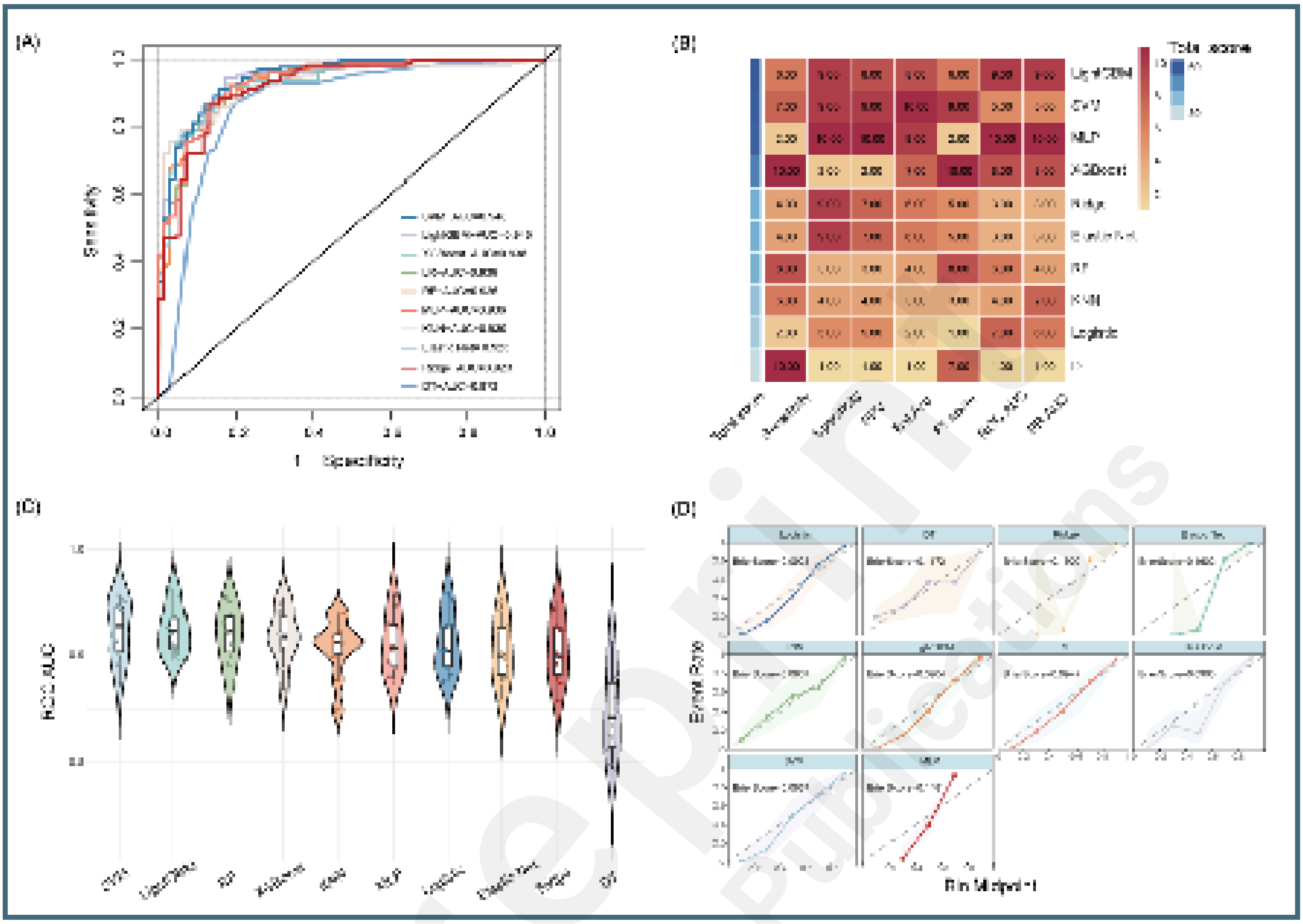
The analysis workflow for the development and evaluation of models.



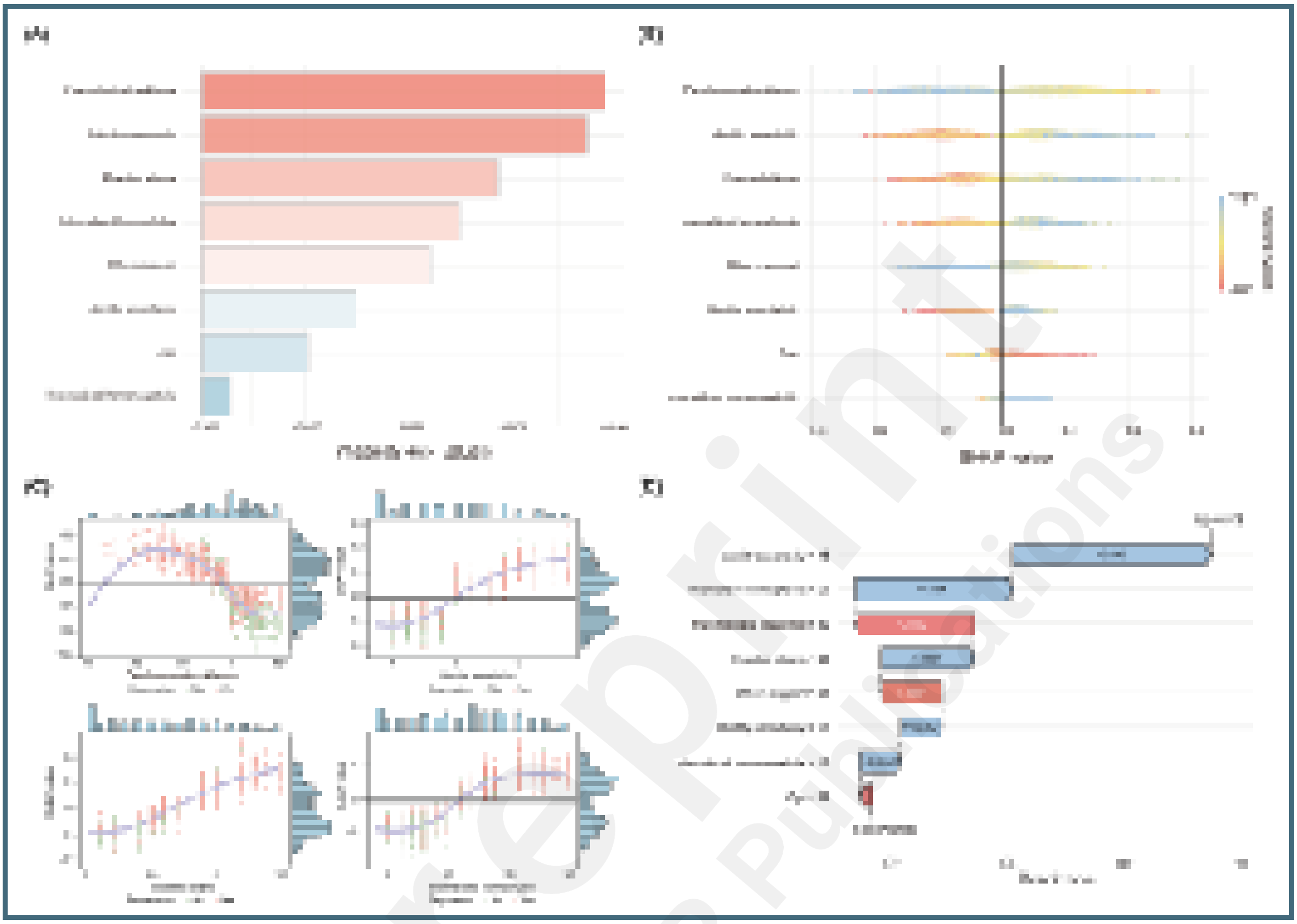
Upset plot of interactions between the predictors.



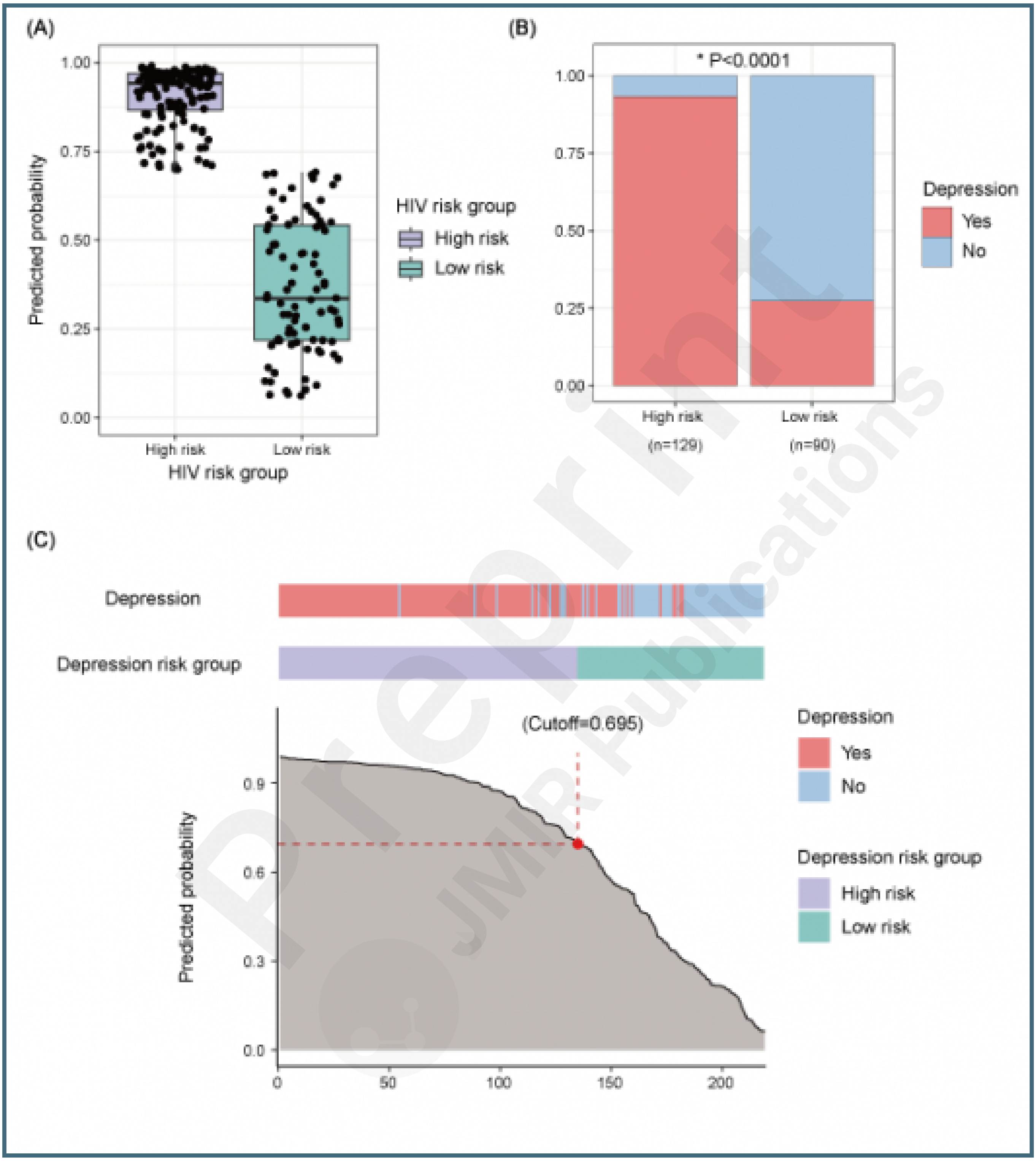
Performance comparison of different machine learning models for depression risk prediction.



SHAP-based interpretation of the optimal machine learning model (LightGBM) for depression risk identification.



Evaluation of the predicted risk scores of the optimal model in the testing set.



Multimedia Appendixes

Supplement of this paper.
URL: http://asset.jmir.pub/assets/6b0ec04d248c70f1018877f1caf8ad04.docx