Original article

Using an interactive web application to identify pre-exposure prophylaxis adherence among men who have sex with men

Bing Lin [a,b], Shihan Feng [a,b], Jiaxiu Liu [c], Kangjie Li [a,b], Guiqian Shi [a,b], Xiaoni Zhong [a,b,*]

[a] School of Public Health, Chongqing Medical University, Chongqing, PR China
[b] Research Center for Medicine and Social Development, Chongqing, PR China
[c] School of Medical Informatics, Chongqing Medical University, Chongqing, PR China

ARTICLE INFO

ABSTRACT

Background: Men who have sex with men (MSM) are at a high risk for HIV infection. While pre-exposure prophylaxis (PrEP) is an effective oral preventive strategy, its success is largely dependent on consistent medication adherence.
Objective: The aim of this study was to develop the machine learning web application and evaluate the performance in predicting PrEP adherence.
Methods: The PrEP prospective cohort study of the MSM population conducted in Western China from 2019 to 2023, and we collected adherence data and personal characteristics data from 747 MSM. Predictor variables were screened and the performance of several machine learning methods in predicting nonadherent behaviors were compared.
Results: A total of 11 candidate variables were screened that predicted nonadherent behaviors. We developed and evaluated five machine learning models that performed well in predicting adherence. Attitudes of male sexual partners, self-efficacy, HIV testing, number of male sexual partners, and risk perception were the most important predictors of adherence. The optimal prediction model was displayed in a shiny web application for online calculation of the probability of occurrence of nonadherent behaviors among MSM.
Conclusions: Machine learning performed well in predicting nonadherent behaviors among MSM. An interactive and intuitive web application can help identify individuals who may have nonadherent behaviors, resulting in improved medication adherence and increased prevention efficacy.

Introduction

Acquired immune deficiency syndrome (AIDS) is a chronic infectious disease caused by the human immunodeficiency virus (HIV). In recent years, the number of men who have sex with men (MSM) who are infected with HIV has shown a rapid growth trend, and MSM have become a high-risk population for HIV infection (Wang et al., 2022). Based on previous reports in the literature, MSM have been found to have a significantly higher likelihood of HIV infection compared to the general population, with some studies indicating it could be around 20 times higher (Bao et al., 2021; Kellie et al., 2018). In China, the prevalence of HIV infection among the MSM population has shown a rising trend over the years, with the proportion of new HIV infections increasing from 2.5 % to 25.5 % (Hu et al., 2019; Qin et al., 2017). To address these challenges and reduce the rate of new HIV infections

among high-risk population, it is important to prioritize HIV/AIDS prevention efforts.

Currently, oral Pre-exposure prophylaxis (PrEP) has been shown to be a safe and effective antiretroviral-based bioprevention approach for HIV (Li et al., 2023). The Chinese Expert Consensus on Drug Use for HIV Pre-Exposure Prophylaxis, with editions released in 2020 and 2023, offers guidance on the use of medications for implementing PrEP (Xu et al., 2020; Zhang et al., 2023). However, PrEP has not been widely promoted in China, and the overall level of PrEP utilization among Chinese MSM is low and the price of the drug is expensive (Jing et al., 2019; Yan et al., 2013). Furthermore, it is worth noting that the efficacy of PrEP is highly dependent on adherence (Claire et al., 2022). Therefore, it is particularly important to conduct studies on PrEP adherence. It not only helps to improve the effectiveness of PrEP, but also provides a scientific basis for the development of more effective interventions and

* Corresponding author at: School of Public Health, Chongqing Medical University, No.1 Medical College Road, Yuzhong District, Chongqing, 400016, PR China.
E-mail address: zhongxiaoni@cqmu.edu.cn (X. Zhong).

policies. Our study on adherence was based on this context, which helped to improve the efficacy of PrEP in the Chinese MSM.

So far, despite the fact that researchers have given PrEP adherence in the MSM population a great deal of attention, the majority of studies have concentrated on exploring the influencing factors of adherence (Andrew et al., 2021; Laura et al., 2021; Luana et al., 2021). Few studies on the prediction of adherence have been reported. Meanwhile, there was no further exploration and excavation in analyzing methods. Given the limitations of PrEP adherence in terms of research content and analytical methods, refining medication-taking strategies and establishing predictive models for identifying nonadherent behaviors could help improve adherence and increase PrEP efficacy. The developed PrEP adherence prediction tool could be a useful clinical service to identify high-risk populations and prioritize resources.

In digital epidemiology, machine learning algorithms are a developing field that may enhance predictive modeling and decision support (Topol, 2019). These methods can incorporate more covariates in large datasets, deal with complex relationships between predictors and outcomes, and achieve high precision (Lee et al., 2018). In China, few studies have used machine learning methods to predict PrEP adherence in the MSM population, and our study will fill this gap. As a result, the applied research on predictive modeling of PrEP adherence was still in the preliminary stage and needed to be better explored and developed. Additionally, the application of predictive modeling is crucial. The ways in which the results of machine learning-based predictions can assist managers or clinicians in a broader range of scenarios deserve significant attention. Mobile health (mHealth) technologies are effective and cost-efficient strategies that can improve individual and public health (Roman et al., 2017). Based on this technology, due to its ubiquity and ease of use on mobile devices, it can be used not only for reminding and recording, but also for expanding application scenarios in combination with modeling (Liu et al., 2020; Roman et al., 2020). Many previous studies have applied machine learning to web applications to enable the prediction of survival, prognosis, and disease progression, which is clinically valuable and significant (Karabacak et al., 2024; Mert et al., 2024; Yan et al., 2024). The user-friendly web application provides an intuitive and easy-to-use platform that makes the application of predictive models more convenient. At the same time, through the Internet, web applications can be accessed anytime and anywhere, increasing their usability and usefulness. The development of an interactive and intuitive web application for predicting adherence levels in MSM population and helping to design individualized interventions would play an important role in improving PrEP adherence.

Therefore, the aim of this study was to develop the PrEP adherence prediction model using different machine learning methods and to develop an interactive and intuitive web application. The resulted machine-learning predictive model for PrEP adherence could be part of an integrated decision-support tool that facilitates the identification of high-risk individuals for nonadherence and personalized prevention strategies in PrEP.

Methods

Study setting, population, and design

Our study was a PrEP prospective cohort study of the MSM population conducted in Western China (Chongqing, Sichuan, and Xinjiang) from 2019 to 2023 (Trial registration: Chinese Clinical Trial ChiCTR190026414). The study population were HIV-negative MSM. The eligible MSM were recruited through collaboration with local Non-governmental organizations (NGOs) and peer referral. Inclusion criteria for study participants included: assigned male sex at birth; age between 18 and 65 years old; HIV-negative; self-reported sexual behavior with a male sexual partner in the last six months; willingness to use PrEP drugs under the supervision of clinicians and to comply with follow-up schedules; and signing of an informed consent form.

The researchers enrolled eligible study participants in the PrEP prospective cohort study and distributed PrEP drugs (Lamivudine and Tenofovir Disoproxil Fumarate tablets) free of charge. Study participants were asked to take PrEP drugs orally once daily, one tablet (300 mg/tablet) at a time. Every 3 months, study participants were informed by the administrator to go to the study center for face-to-face follow-up and a new round of PrEP drugs were dispensed. A total of 4 follow-up visits were conducted, at months 3, 6, 9, and 12.

Outcomes

Study participants self-reported their adherence each time they attended a follow-up visit. Study participants were asked, "In the last two weeks, have you had a missed dose? How many days were missed?" Participants answered "yes" or "no" and the number of days missed (0–14 days). Adherence is equal to the percentage of days on medication. Previous literature suggested that adherence of $\geq 80\%$ could be defined as good adherence (David et al., 2018; Haberer et al., 2013; Ira et al., 2016).

Predictors

In total, we measured 31 variables of MSM personal characteristics (Supplementary Table 1). After enrollment in the PrEP cohort, study participants were randomized into the reminder group and the no-reminder group. Participants in the reminder group received daily reminder messages based on the mobile app; participants in the no-reminder group who took PrEP drug daily did not receive reminder messages. Reminders were sent through the WeChat App (one of the most popular social media apps in China) at regular intervals to remind users to take their PrEP drugs. We named this variable "reminder system" to explore whether the use of the reminder system as a potential predictor of adherence was effective.

The main HIV-related characteristics of MSM individuals included HIV knowledge score, HIV testing and counseling, risk perception, and number of male sexual partners. Of these, the HIV knowledge score was measured by the knowledge scale consisting of 13 questions related to HIV (Diclemente et al., 1986; Koopman et al., 1990). These questions included some assessments of HIV knowledge, where participants answered "True", "False", or "Don't know". A correct answer was scored as 1 point, while an incorrect answer or "Don't know" was scored as 0 points, resulting in a total scale score ranging from 0 to 13. A score of ≥ 11 indicated a high level of HIV knowledge (Qu et al., 2018).

Variables related to PrEP primarily included attitude of male sexual partners and self-efficacy. Attitude of male sexual partners were measured by a question, "How do you think your male partners would react if they knew you were taking PrEP?" Participants chose to categorize the variables by answering "positive", "neutral", and "negative". MSM were asked about "I am confident that I can remember to take my medicine on time" as a measure of their self-efficacy and were categorized into high and low levels.

Model development and validation

Data on adherence at the first follow-up (month 3) were taken as the outcome variable ($\geq 80\%$ was defined as high adherence), and data on personal characteristics of MSM at baseline were taken as predictor variables. We used five prediction methods, including a traditional Logistic regression model (LR) and four machine learning algorithms. The machine learning algorithms were Decision tree (DT), Support vector machine (SVM), Random forest (RF), and Extreme gradient boosting machine (XGBoost). The data were preprocessed before being analyzed. First, the multiple imputation method was used to fill in the missing values of each predictor variable, which was performed using SAS software. Secondly, LASSO regression was used to screen for potential predictors and to make a preliminary comparison of the differences

between the screened variables in the high and low adherence groups. This step was performed using the "glmnet" package in the R software.

We split the data into the 80 % training set and the 20 % testing set. The SMOTE algorithm was used to balance the data, and the balanced training data were brought into the five prediction models for training and hyperparameter optimization. The modeling results with the best performance were saved, applied to the testing set data, and cross-validated with ten folds. The metrics for model evaluation included Area under the curve of receiver operating characteristic (AUC-ROC), Accuracy, Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) and F1-Score. Our study used AUC-ROC as the primary evaluation metric and Sensitivity as the secondary evaluation metric. The DeLong test was used to compare the differences between the AUC-ROCs to find the model with the best predictive performance. The training and testing processes of the models were realized by using the "tidymodels" package in R software.

The trained optimal model was applied to 6-month, 9-month, and 12-month data as an internal validation to assess the predictive effectiveness of the model. Meanwhile, for the optimal model, the importance of its variables was ranked and the partial dependence plots of the top 5 variables were created to explain the effects of the variables. The predictions of the optimal model were also displayed in a shiny web application, which enabled online probability estimation and interactive computation. This step was realized using the "shiny" package in R software. The flowchart of the study was shown in Fig. 1.

Sensitivity analysis

In the optimal model, we trained the model by bringing each predictor variable into the model and compared the resulting univariate AUC-ROC with the AUC-ROC of the optimal model. Secondly, for the subgroups of adherence, we also set the criteria to 90 %, 100 %, and 60 %, respectively, and brought them into the optimal model. In addition, previous findings have shown that men who have sex with multiple men (MSMM) have a higher risk of HIV infection than other men (Jiaxiu et al., 2021). We categorized MSM with two or more male sexual partners as MSMM and the rest as Non-MSMM. On the one hand, both MSMM and Non-MSMM were applied to the optimal model to evaluate the prediction effect; on the other hand, separate prediction models were built for MSMM and Non-MSMM to evaluate the prediction effect. The performance of the two prediction methods in MSMM and Non-MSMM was compared.

Results

Description of demographic information

A total of 747 MSM participated in the first follow-up, and their basic demographic information were described in Supplementary Table 2. According to the results of the descriptive analysis, most of the recruited MSM were over 25 years old, mainly urban population, Han ethnicity,
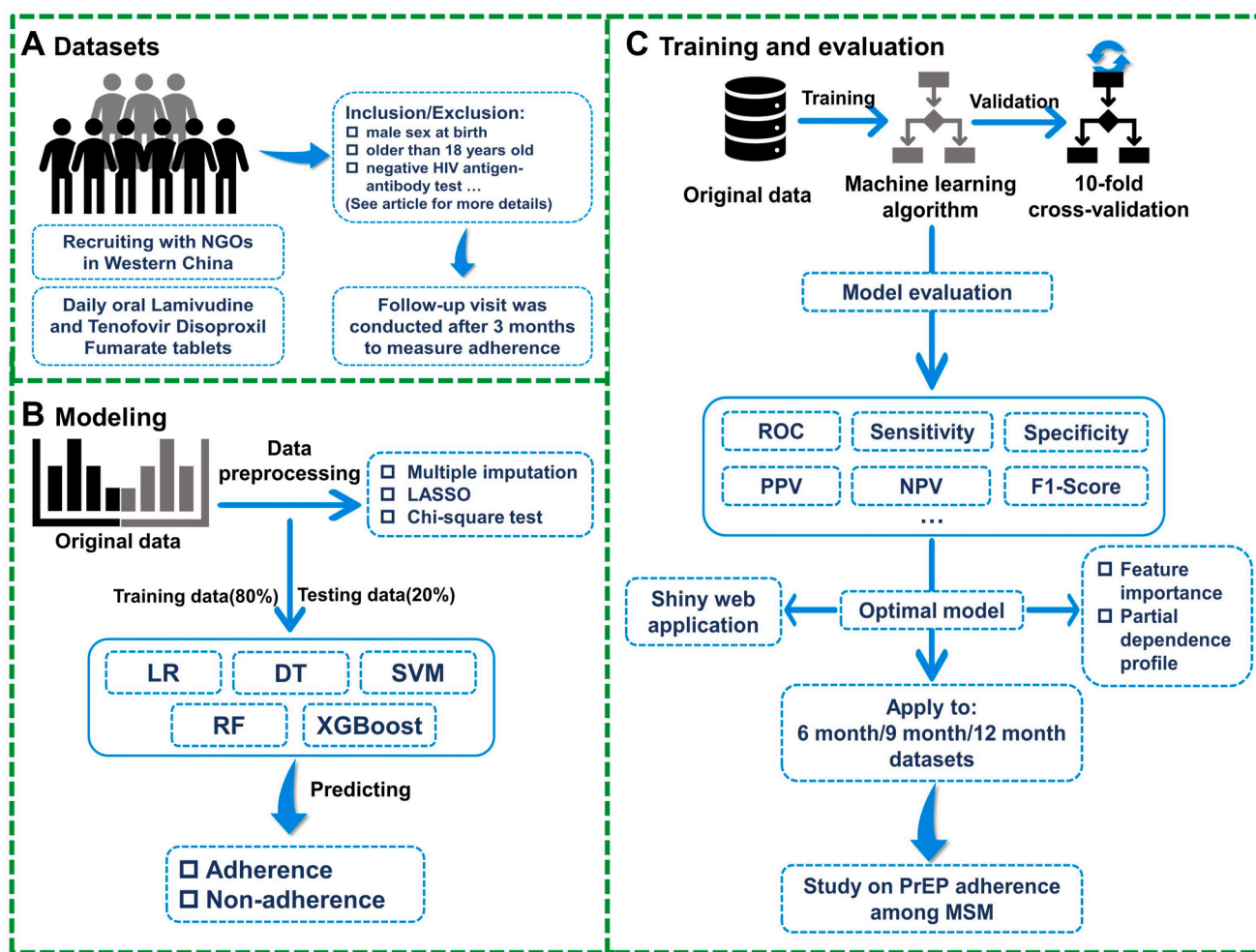

Fig. 1. Research flowchart

NGOs: Non-governmental organizations; LASSO: Least absolute shrinkage and selection operator; LR: Logistic regression; DT: Decision tree; SVM: Support vector machine; RF: Random forest; XGBoost: Extreme gradient boosting machine; ROC: Receiver operating characteristic curve; PPV: Positive predictive value; NPV: Negative predictive value; PrEP: Pre-exposure prophylaxis; MSM: Men who have sex with men.

and highly educated. 57.56 % of MSM had high adherence.

Variable screening

The process and results of the LASSO screening were shown in Fig. 2. According to the results presented in Fig. 2, a total of 11 variables were filtered out.

Univariate analysis for different adherence groups

The 11 variables screened were compared between the two groups (Supplementary Table 3). It was observed that the variable "Reminder system" was not statistically different between the high and low adherence groups ($p = 0.065$), while the rest of the variables showed significant differences between the two groups.

Model construction

The screened 11 variables were brought into the model as independent variables, split into the training set (80 %) and the testing set (20 %), and balanced using the SMOTE algorithm (Supplementary Table 4 for the comparison of the training and testing sets). LR, DT, SVM, RF and XGBoost methods were used to train and predict the model (setting low adherence as a positive result, in which the predicted result was the nonadherent MSM). The evaluation metrics of each prediction model on the testing set were shown in Table 1. According to the prediction results of the model, RF performed best on the ROC-AUC metric, with a value of 0.704 (95 % CI=0.627–0.781); RF performed best on the Sensitivity metric, with a value of 67.44 %.

The results of the AUC-ROC comparison of the five prediction models were shown in Table 2. It was observed that the AUC-ROC of the RF prediction model (AUC-ROC=0.704, 95 %CI=0.627–0.781)
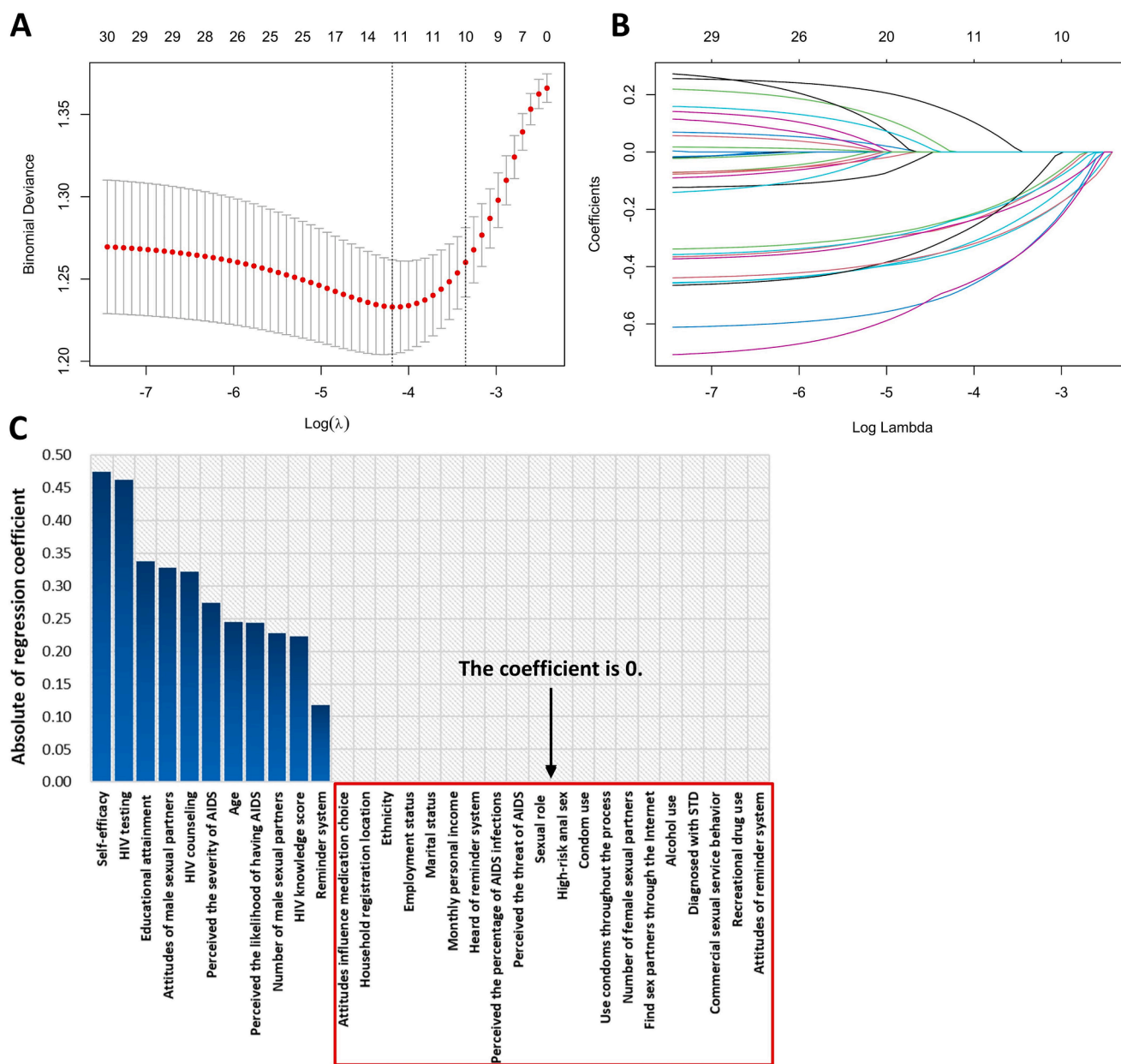

Fig. 2. The processes and results of LASSO screening variables
LASSO: Least absolute shrinkage and selection operator; HIV: Human immunodeficiency virus; AIDS: Acquired immune deficiency syndrome; STD: Sexually transmitted disease.

Table 1

Evaluation metrics for each predictive model on the testing set.

Models	AUC-ROC (95 %CI)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1-Score
LR	0.660 (0.579–0.741)	0.6105	0.6047	0.6163	0.6118	0.6092	0.6082
DT	0.656 (0.578–0.734)	0.6163	0.6163	0.6163	0.6163	0.6163	0.6163
SVM	0.661 (0.580–0.742)	0.6105	0.5116	0.7093	0.6377	0.5922	0.5677
RF	0.704 (0.627–0.781)	0.6337	0.6744	0.5930	0.6237	0.6456	0.6480
XGBoost	0.675 (0.595–0.755)	0.5930	0.5814	0.6047	0.5952	0.5909	0.5882

LR: Logistic regression; DT: Decision tree; SVM: Support vector machine; RF: Random forest; XGBoost: Extreme gradient boosting machine; AUC-ROC: Area under the curve of receiver operating characteristic; PPV: Positive predictive value; NPV: Negative predictive value.

Table 2

Results of AUC-ROC comparison of five predictive models.

Models	LR	DT	SVM	RF	XGBoost
LR	–	0.917	0.983	<0.05	0.492
DT	0.917	–	0.924	0.142	0.587
SVM	0.983	0.924	–	0.248	0.723
RF	<0.05	0.142	0.248	–	<0.05
XGBoost	0.492	0.587	0.723	<0.05	–

LR: Logistic regression; DT: Decision tree; SVM: Support vector machine; RF: Random forest; XGBoost: Extreme gradient boosting machine; AUC-ROC: Area under the curve of receiver operating characteristic.

outperformed the traditional Logistic regression model (AUC-ROC=0.660, 95 %CI=0.579–0.741), and the XGBoost model (AUC-ROC=0.675, 95 %CI=0.595–0.755).

Model evaluation and feature importance

Fig. 3-A illustrated the AUC-ROC results for the five predictive model. RF performed optimally in predicting PrEP adherence in the MSM population. Fig. 3-B illustrated the results of the five prediction models in the ten-fold cross-validation. All five prediction models showed stability in the ten-fold cross validation. The 6-month, 9-month, and 12-month follow-up data were brought into the trained RF model (Fig. 3-C). The ranked order of the variables for the optimal model was shown in Fig. 3-D. Based on the partial dependence plots of top 5 most important variables (Fig. 3-E), MSM with negative attitudes of male sexual partners, low self-efficacy, not having been tested for HIV, not having a male sexual partner, and having low risk perception were more likely to engage in nonadherent behaviors.

Web application

The predictions of the optimal model were eventually displayed in a shiny web application for online estimation of probabilities and interactive computation (Fig. 4), which was available at: https://lauralin1104.shinyapps.io/RFshiny/.

Sensitivity analysis

Based on the sensitivity analyses, we obtained that (1) the inclusion of a single variable as a predictor variable in the predictive model was not as good as the full-variable model (Supplementary Table 5). (2) The optimal model performed well in predicting different adherence categorization criteria (Supplementary Fig. 1). (3) The optimal model could be used to directly predict adherence in different populations (Supplementary Table 6).

Discussion

Our findings indicated that the machine learning approaches were accurate in predicting PrEP adherence. The most useful predictors were routine and easy to collect. The results of the sensitivity analyses showed that the trained optimal model could be used in different categorization

criteria for adherence, as well as applied to different populations of MSM. The routine collection of these data and the accuracy of their predictive results suggested that machine learning approaches may be of value in routine PrEP medication management services. In addition, machine learning predictive modeling was applied to the web application to help identify nonadherent individuals in the MSM population, which can help improve medication adherence and increase prevention efficacy in high-risk population.

In our study, the MSM in Western China had an adherence rate of 57.56 % in the use of PrEP. Although this rate is relatively high, it is still not sufficient to ensure optimal protection against HIV, highlighting the need for further interventions to improve adherence. Therefore, timely identification and intervention are necessary for MSM who were nonadherent during PrEP. Adherence of study participants was measured using self-report. Previous studies have shown that this method may result in inaccurate estimation of adherence due to recall bias and social desirability bias (Buscher et al., 2011; Williams et al., 2013). However, it was also noted in the literature that self-reported data from the participants were reliable (Marins et al., 2019). The previous study of the PrEP program from a clinical setting, researchers found a significant correlation between self-reported adherence and medication levels, demonstrating the feasibility and validity of self-reports for measuring PrEP adherence (Montgomery et al., 2017). Compared to other methods of measuring adherence, such as measuring Tenofovir-diphosphate (TFV-DP) concentration in dried blood spots (DBSs) (Goldwirt et al., 2021), electronic pill cap (Atukunda et al., 2022), drug level monitoring in hair and urine (Hannaford et al., 2021), self-reported adherence was more convenient and quicker to use for initial assessment of early clinical adherence and was also more easily available in any setting. Meanwhile, compared with the classical HIV/STIs prediction model, the AUC values of the individual models for predicting HIV ranged from 0.698 to 0.763 (Bao et al., 2021). The mean value of the AUC of our RF prediction model for adherence in ten-fold cross-validation was 0.757, indicating that RF performed well and reliably in predicting adherence in the MSM population. It was evident that the predictive model of PrEP adherence that we developed using MSM individual characteristic data, with variables that were easy to collect in daily routine, was feasible and valuable in practical applications.

Our findings indicated that attitudes of male sexual partners were a significant variable in predicting PrEP adherence in the MSM population and ranked first. MSM whose male sexual partners' attitudes were negative were more likely to engage in nonadherent behaviors. This was an interesting finding because few studies in China have related the attitudes of male sexual partners to PrEP adherence in the MSM population. In a study conducted with same-sex male couples, 26.7 % reported that their sexual partners did not support the use of PrEP, and 57.7 % did not believe that PrEP would be beneficial to them or their sexual partners (Kahle et al., 2020). However, in China, influenced by traditional culture, it is very difficult to conduct surveys among male couples. The cultural and social environment in China has low acceptance of MSM, and many MSM individuals face significant pressure and stigma from their families, society, and workplaces (Zhou et al., 2024). This environment may lead MSM individuals to have more conservative and resistant attitudes towards openly discussing and adopting sexual
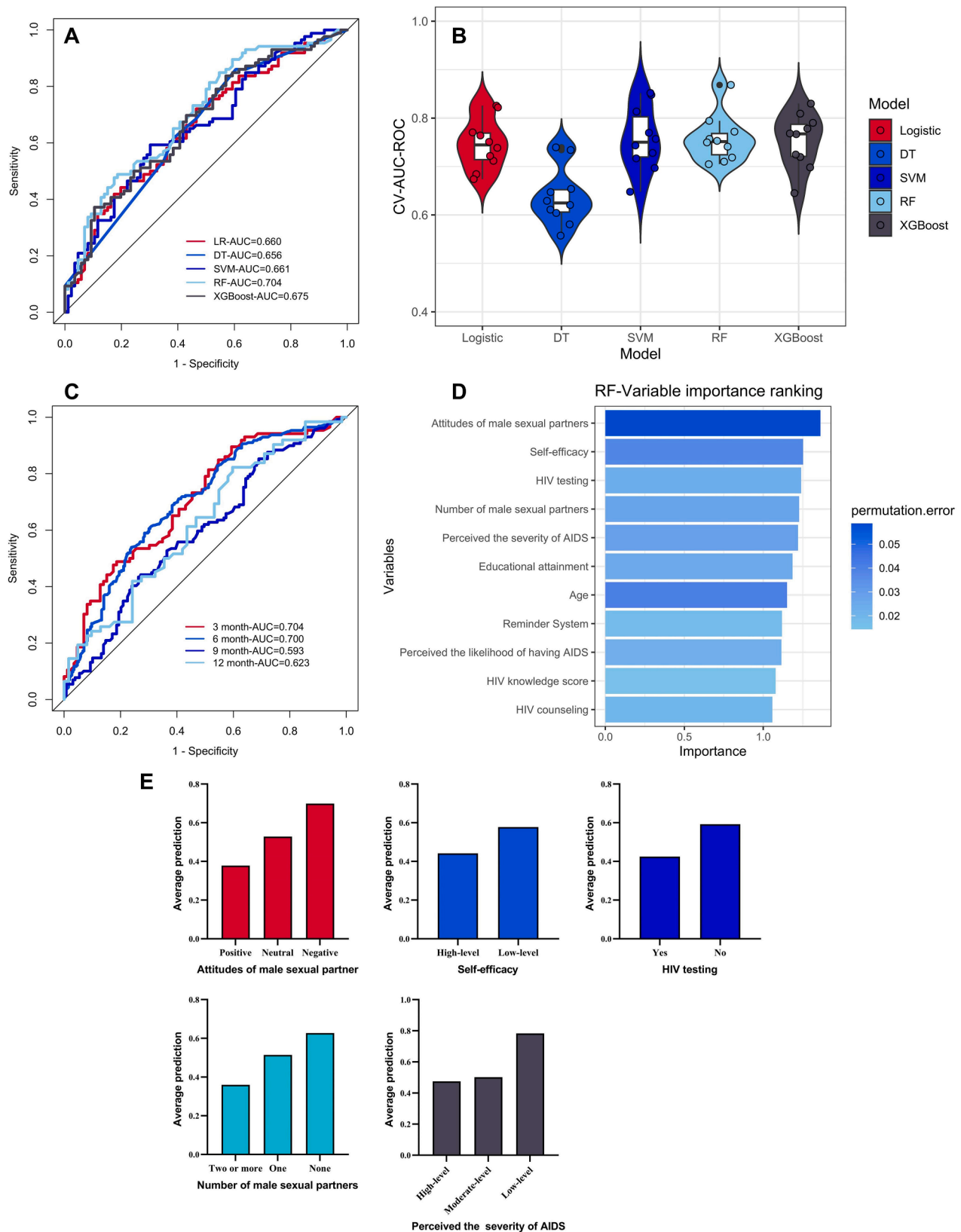
Fig. 3. Summary of model prediction results, variable importance and partial dependence plot. A AUC-ROC results for five prediction models. B Results of the five predictive models in ten-fold cross-validation. C Results of bringing 6-month, 9-month, and 12-month follow-up data to the optimal model. D Ranking of variable importance for the optimal model. 3-E Partial dependence plot of the top 5 ranked variables.

LR: Logistic regression; DT: Decision tree; SVM: Support vector machine; RF: Random forest; XGBoost: Extreme gradient boosting machine; CV: Cross validation; AUC: Area under the curve; ROC: Receiver operating characteristic curve; HIV: Human immunodeficiency virus; AIDS: Acquired immune deficiency syndrome.

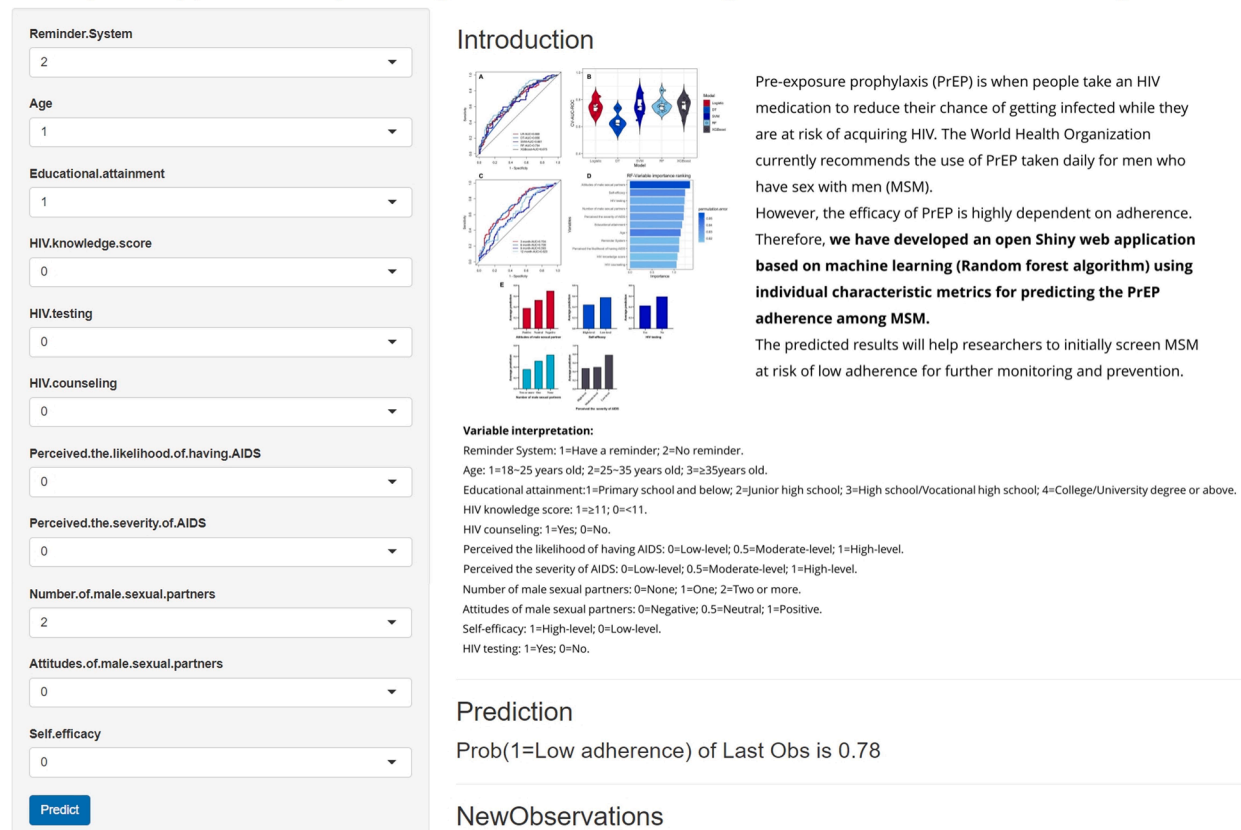**Fig. 4.** A Shiny web application for predicting PrEP adherence among MSM based on machine learning (https://lauralin1104.shinyapps.io/RFshiny/).
HIV: Human immunodeficiency virus; AIDS: Acquired immune deficiency syndrome; PrEP: Pre-exposure prophylaxis; MSM: Men who have sex with men.

behavior-related preventive measures, such as PrEP. Previous research showed that the use of peer education intervention can encourage MSM to access HIV prevention information, counseling, and testing (Duan et al., 2013). Perceived peer use of PrEP was positively correlated with personal use of PrEP (Mike et al., 2020). It was concluded that MSM's male sexual partners, as the manifestation of an intimate peer relationship, were likely to influence their interest, willingness, and adherence to PrEP, which highlighted the need for tailored dyadic (users and male sexual partners) interventions. At the same time, future study should consider the impact of cultural background on the health behaviors and attitudes of MSM individuals. Cross-cultural comparative studies can be conducted to further explore the awareness and acceptance of PrEP among MSM population in different cultural contexts.

Self-efficacy in MSM population was also one of the important predictors of their adherence behavior. Previous studies mentioned that adherence self-efficacy was associated with PrEP knowledge and confidence in HIV prevention effectiveness (Sun et al., 2021). In addition, previous study in Mississippi Southern MSM showed that greater self-efficacy for taking PrEP was more likely to be adherent to PrEP (Whiteley et al., 2021). This was consistent with our findings. At the same time, HIV testing as one of the predictive variables for adherence suggested that we need to focus on the population who have not been tested for HIV, and these MSM may be at a higher risk of infection. HIV testing and PrEP were potentially complementary tools that could be combined to enhance the ability of at-risk individuals to prevent HIV infection (Mujugira et al., 2022). In our study, the number of male sexual partners and risk perception were shown to be significantly associated with PrEP adherence, which was also consistent with previous research findings (Lebouché et al., 2016; Wood et al., 2019). Our study used the variable reminder system to predict adherence in the

MSM population. Previous studies have found that forgetting to take medication was the most common objective cause of nonadherence, accounting for 70.21 % (Qu et al., 2018). Our findings suggested that APP-based daily reminder mechanisms were associated with adherence in the MSM population. Furthermore, the effectiveness of mobile health (mHealth) intervention in promoting adherence has been observed in various populations. For instance, previous studies have found that digital health interventions in relation to obesity management have the potential to achieve relatively high rates of adherence (Jakob et al., 2024). Meanwhile, mHealth can also be used to support self-management of cancer pain (Wu et al., 2024). It highlights the importance of mHealth, with customizable reminders, tailored content, and personalized services that will help increase the effectiveness of digital health interventions in promoting healthy behaviors.

Strengths and limitations

All predictors in our model were self-reported variables, so our predictive tools may have real-life applications. Based on the shiny web application we developed, a user-friendly, internet-based self-assessment tool could potentially help the MSM population identify nonadherent behaviors during PrEP use in a clinical setting. This allowed clinicians or administrators to identify individuals at high risk for nonadherent behaviors and provide timely intervention and medication management. In the clinical setting, we believe that this tool would help clinicians categorize high-risk individuals and guide the use of PrEP more effectively. By being more alert to individuals with potential nonadherent behaviors, clinicians can offer personalized support and interventions, such as more frequent follow-ups, counseling, and reminders. In addition, this tool could be integrated into electronic

health records (EHR) systems to streamline the process of identifying and monitoring high-risk individuals. For example, alerts could be generated for healthcare providers when a patient is flagged as high-risk, prompting immediate action. Moreover, risk assessment of individuals provides quantitative information about the overall risk of PrEP nonadherence behaviors in the MSM population, which is crucial for behavioral and epidemiologic surveillance. This data can inform public health strategies and policies aimed at improving adherence rates, such as targeted education campaigns and community-based support programs.

There were some limitations to our study. Firstly, as a sensitive population, MSM attach great importance to privacy issues. Since the investigation in this study was in the form of a face-to-face questionnaire and all variables were self-reported information, some sensitive issues may not be authentically represented. More emphasis should be placed on privacy protection for sensitive populations in future investigations. Secondly, our study data sources were subjective and lacked objective measures. For example, measures of adherence lacked objective data. However, for objective measures of adherence, for instance, measuring blood levels in the user's body, which were invasive, costly, and time-consuming, were not necessarily widely accepted and used on a large scale by the MSM in the clinic. Self-reported methods had the advantage of being convenient and quick to use and could be widely used in the clinical setting as a measurement tool for initial screening. In addition, during further validation of the model, we found that the predictive efficacy of the optimal model performs well when 6 months of data were brought in. However, the predictive efficacy decreased with the 9 and 12 months of data. We analyzed that this may be related to loss to follow-up (LFU) among MSM in the late follow-up period. Previous studies have also shown a higher risk of PrEP LFU among younger and less educated MSM (Amanda et al., 2022). These variables were disproportionately lost during follow-up and may have affected the predictive efficacy of adherence. Currently, our application is designed for online visualization only, and we are actively considering implementing more functions in a future update to enhance the usability of our application. In the future, if the opportunity arises, we can investigate the effects of the model's predictions on other populations, thereby extending the utility of the web application. Finally, our study used only a few of the most commonly used machine learning algorithms for adherence prediction, and other methods could be explored in future research. Since the model was constructed only in the Western region of China, further exploration and validation were needed to extend the optimal model to the whole country to make it universally applicable. Meanwhile, fuller external validation should be conducted in the future.

Conclusion

In conclusion, our study demonstrated that machine learning algorithms excel in predicting nonadherent behaviors among MSM using PrEP. This innovative online assessment tool represents a substantial advancement over traditional methods by offering real-time, personalized risk assessments and interventions. It stands out by leveraging predictive analytics to proactively address nonadherence, which is crucial for improving health outcomes in a vulnerable population. The tool's ability to dynamically adapt and learn from new data positions it as an invaluable asset in the public health arsenal, particularly in the monitoring and management of PrEP adherence among MSM in Western China. Future research should explore expanding this tool to other high-risk populations to enhance its generalizability and public health impact. In addition, integrating new predictive variables, such as psychosocial factors and behavioral patterns, could further refine its accuracy and utility. This continuous development and application of machine learning technologies hold promise for advancing public health strategies and optimizing PrEP adherence.

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.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijchp.2024.100490.

References

Amanda, G., Lara, C., Valdiléa, V., Maria, P., Brenda, H., Ronaldo, M., et al. (2022). Travestis, transgender women and young MSM are at high risk for PrEP early loss to follow-up in Rio de Janeiro, Brazil. In *The Brazilian journal of infectious diseases: An official publication of the Brazilian Society of Infectious Diseases, 27,* Article 102733.

Andrew, G., Luke, P., R, C. J., Shanta, Z., Nancy, M., & Katherine, F (2021). Factors associated with retention and adherence in a comprehensive, diverse HIV Pre-Exposure Prophylaxis (PrEP) clinic. *AIDS Research and Human Retroviruses, 38*(4), 327–335.

Atukunda, E., Owembabazi, M., Pratt, M., Psaros, C., Muyindike, W., Chitneni, P., et al. (2022). A qualitative exploration to understand barriers and facilitators to daily oral PrEP uptake and sustained adherence among HIV-negative women planning for or with pregnancy in rural Southwestern Uganda. *Journal of the International AIDS Society, 25*(3), e25894.

Bao, Y., Medland, N., Fairley, C., Wu, J., Shang, X., Chow, E., et al. (2021). Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. *The Journal of Infection, 82*(1), 48–59.

Buscher, A., Hartman, C., Kallen, M., & Giordano, T. (2011). Validity of self-report measures in assessing antiretroviral adherence of newly diagnosed, HAART-naïve, HIV patients. *HIV Clinical Trials, 12*(5), 244–254.

Claire, B., Adiba, H., David, M., Michael, D., Katya, C., Karen, C., et al. (2022). Sexually transmitted infections among participants in an HIV PrEP adherence trial. *International Journal of STD & AIDS, 33*(4), 397–403.

David, S., Mia, E. A., & Susanne, S. (2018). A systematic review of adherence to oral pre-exposure prophylaxis for HIV - how can we improve uptake and adherence? *BMC Infectious Diseases, 18*(1), 581.

Diclemente, R., Jim, Z., & Lydia, T. (1986). Adolescents and AIDS: A survey of knowledge, attitudes and beliefs about AIDS in San Francisco. *American Journal of Public Health, 76*(12), 1443.

Duan, Y., Zhang, H., Wang, J., Wei, S., Yu, F., & She, M. (2013). Community-based peer intervention to reduce HIV risk among men who have sex with men in Sichuan province, China. *AIDS Education and Prevention: Official Publication of the International Society for AIDS Education, 25*(1), 38–48.

Goldwirt, L., Bauer, R., Liegeon, G., Charreau, I., Delaugerre, C., Cotte, L., et al. (2021). Estimated pill intake with on-demand PrEP with oral TDF/FTC using TFV-DP concentration in dried blood spots in the ANRS IPERGAY trial. *The Journal of Antimicrobial Chemotherapy, 76*(10), 2675–2680.

Haberer, J., Baeten, J., Campbell, J., Wangisi, J., Katabira, E., Ronald, A., et al. (2013). Adherence to antiretroviral prophylaxis for HIV prevention: A substudy cohort within a clinical trial of serodiscordant couples in East Africa. *PLoS Medicine, 10*(9), Article artnoe1001511.

Hannaford, A., Arens, Y., & Koenig, H. (2021). Real-time monitoring and point-of-care testing: A review of the current landscape of PrEP adherence monitoring. *Patient Preference and Adherence, 15,* 259–269.

Hu, Y., Zhong, X., Peng, B., Zhang, Y., Liang, H., Dai, J., et al. (2019). Comparison of depression and anxiety between HIV-negative men who have sex with men and women (MSMW) and men who have sex with men only (MSMO): A cross-sectional study in Western China. *BMJ Open, 9*(1), Article e023498.

Ira, W., Yoojin, L., Joanne, M., Floyd, F., & William, R. (2016). Validation of a new three-item self-report measure for medication adherence. *AIDS and Behavior, 20*(11), 2700–2708.

Jakob, R., Narauskas, J., Fleisch, E., König, L. M., & Kowatsch, T. (2024). Factors associated with adherence to a public mobile nutritional health intervention: Retrospective cohort study. *Computers in Human Behavior Reports, 15*, Article 100445.

Jiaxiu, L., Ruibin, D., Bing, L., Hong, P., Yuwen, G., Jianghong, D., et al. (2021). Risk management on pre-exposure prophylaxis adherence of men who have sex with multiple men: a multicenter prospective cohort study. *Risk Management and Healthcare Policy, 14*, 1749–1761.

Jing, H., Jennifer, B., Liming, W., Guodong, M., Zihuang, C., Ying, H., et al. (2019). PrEP uptake preferences among men who have sex with men in China: Results from a National Internet Survey. *Journal of the International AIDS Society, 22*(2), e25242.

Kahle, E., Sharma, A., Sullivan, S., & Stephenson, R. (2020). The influence of relationship dynamics and sexual agreements on perceived partner support and benefit of PrEP use among same-sex male couples in the U.S. *AIDS and Behavior, 24*(7), 2169–2177.

Karabacak, M., Carr, M. T., Schupper, A. J., Bhimani, A. D., Steinberger, J., & Margetis, K. (2024). An interpretable machine learning approach to predict survival outcomes in spinal and sacropelvic chordomas. *Spine, 10*, 1097.

Kellie, F., & Carmen, P. (2018). Does pre-exposure prophylaxis for HIV prevention in men who have sex with men change risk behaviour? A systematic review. *Journal of Clinical Nursing, 27*(17–18), 3254–3265.

Koopman, Borus, Henderson, R., & Hunter, Bradley (1990). Assessment of knowledge of AIDS and beliefs about AIDS prevention among adolescents. *AIDS education and prevention: Official Publication of the International Society for AIDS Education, 2*(1), 58–69.

Laura, W., Lacey, C., Shufang, S., Nicholas, T., Dylan, H., Jesse, M., et al. (2021). Factors associated with PrEP adherence among MSM living in Jackson, Mississippi. *Journal of HIV/AIDS & Social Services, 20*(3), 246–261.

Lebouché, B., Engler, K., Machouf, N., Lessard, D., & Thomas, R. (2016). Predictors of interest in taking pre-exposure prophylaxis among men who have sex with men who used a rapid HIV-testing site in Montreal (Actuel sur Rue). *HIV Medicine, 17*(2), 152–158.

Lee, H.-C., Yoon, S. B., Yang, S.-M., Kim, W. H., Ryu, H.-G., Jung, C.-W., et al. (2018). Prediction of acute kidney injury after liver transplantation: Machine learning approaches vs. logistic regression model. *Journal of Clinical Medicine, 7*(11), 428.

Li, C., Xiong, Y., Liu, H., Luo, D., Tucker, J., Maman, S., et al. (2023). Multifaceted barriers to rapid roll-out of HIV pre-exposure prophylaxis in China: A qualitative study among men who have sex with men. *International Journal of Behavioral Medicine*, 1–11.

Liu, A., Laborde, N., Coleman, K., Vittinghoff, E., Gonzalez, R., Wilde, G., et al. (2020). DOT Diary: Developing a novel mobile app using artificial intelligence and an electronic sexual diary to measure and support PrEP adherence among young men who have sex with men. *AIDS and Behavior, 25*(4), 1–12.

Luana, M. S. M., Thiago, S. T., Mendes, L. P., I, M. R., C, L. I., Brenda, H., et al. (2021). Factors associated with self-reported adherence to daily oral pre-exposure prophylaxis among men who have sex with man and transgender women: PrEP Brasil study. *International Journal of STD & AIDS, 32*(13), 1231–1241.

Marins, L., Torres, T., Leite, I., Moreira, R., Luz, P., Hoagland, B., et al. (2019). Performance of HIV pre-exposure prophylaxis indirect adherence measures among men who have sex with men and transgender women: Results from the PrEP Brasil Study. *PloS One, 14*(8), Article e0221281.

Mert, K., Abhiraj, B., Alexander, S., Matthew, C., Jeremy, S., & Konstantinos, M. (2024). Machine learning models on a web application to predict short-term postoperative outcomes following anterior cervical discectomy and fusion. *BMC Musculoskeletal Disorders, 25*(1), 401. -401.

Mike, P., Nathaniel, W., Frank, D., & Keenan, P. (2020). Psychosocial barriers to pre-exposure prophylaxis (PrEP) uptake: The roles of heterosexual self-presentation, sexual risk, and perceived peer prep use. *Psychology of Men & Masculinities, 21*(4), 699–703.

Montgomery, M. C., Oldenburg, C. E., Nunn, A. S., Mena, L., Anderson, P., Liegler, T., et al. (2017). Adherence to pre-exposure prophylaxis for HIV prevention in a clinical setting. *PloS One, 11*(6), Article e0157742.

Mujugira, A., Nakyanzi, A., Nabaggala, M., Muwonge, T., Ssebuliba, T., Bagaya, M., et al. (2022). Effect of HIV self-testing on PrEP adherence among gender-diverse sex workers in Uganda: A randomized trial. *Journal of Acquired Immune Deficiency Syndromes (1999), 89*(4), 381–389.

Qin, Q., Guo, W., Tang, W., Mahapatra, T., Wang, L., Zhang, N., et al. (2017). Spatial analysis of the human immunodeficiency virus epidemic among men who have sex with men in China, 2006-2015. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, 64*(7), 956–963.

Qu, D., Zhong, X., Xiao, G., Dai, J., Liang, H., & Huang, A. (2018). Adherence to pre-exposure prophylaxis among men who have sex with men: A prospective cohort study. *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases, 75*, 52–59.

Roman, S., Frederick, A., Elizabeth, D., Brian, S., Yerina, R., & Michael, C. (2020). Feasibility and acceptability of an mHealth-based approach as an HIV prevention strategy among people who use drugs on pre-exposure prophylaxis. *Patient Preference and Adherence, 14*, 107–118.

Roman, S., Pramila, K., & Michael, C. (2017). Interest in use of mHealth technology in HIV prevention and associated factors among high-risk drug users enrolled in methadone maintenance program. *AIDS Care, 29*(9), 1144–1148.

Sun, S., Yang, C., Zaller, N., Zhang, Z., Zhang, H., & Operario, D. (2021). PrEP willingness and adherence self-efficacy among men who have sex with men with recent condomless anal sex in urban China. *AIDS and Behavior, 25*(11), 3482–3493.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56.

Wang, H., Wang, Z., Huang, X., Chen, Y., Wang, H., Cui, S., et al. (2022). Association of HIV preexposure prophylaxis use with HIV incidence among men who have sex with men in China: A nonrandomized controlled trial. *JAMA Network Open, 5*(2), Article e2148782.

Whiteley, L., Craker, L., Sun, S., Tarantino, N., Hershkowitz, D., Moskowitz, J., et al. (2021). Factors associated with PrEP adherence among MSM living in Jackson, Mississippi. *Journal of HIV/AIDS & Social Services, 20*(3), 246–261.

Williams, A., Amico, K., Bova, C., & Womack, J. (2013). A proposal for quality standards for measuring medication adherence in research. *AIDS and Behavior, 17*(1), 284–297.

Wood, S., Gross, R., Shea, J., Bauermeister, J., Franklin, J., Petsis, D., et al. (2019). Barriers and facilitators of PrEP adherence for young men and transgender women of color. *AIDS and Behavior, 23*(10), 2719–2729.

Wu, W., Graziano, T., Salner, A., Chen, M. H., Judge, M. P., Cong, X., et al. (2024). Acceptability, effectiveness, and roles of mHealth applications in supporting cancer pain self-management: integrative review. *JMIR mHealth and uHealth, 12*, e53652.

Xu, J., Huang, X., Liu, X., Wang, L., Chen, Y., Wang, H., et al. (2020). Chinese expert consensus on HIV pre-exposure prophylaxis. *Chinese Journal of AIDS & STD, 26*(11), 1265–1271.

Yan, Z., Bin, P., Ying, S., Hao, L., Hong-Bin, P., Han-Zhu, Q., et al. (2013). Attitudes toward HIV pre-exposure prophylaxis among men who have sex with men in western China. *AIDS Patient care and STDs, 27*(3), 137–141.

Yan, Z., Shi, Z., Zhu, Q., Feng, J., Liu, Y., Li, Y., et al. (2024). Interpretable and intuitive machine learning approaches for predicting disability progression in relapsing-remitting multiple sclerosis based on clinical and gray matter atrophy indicators. *Academic Radiology, 31*(7), 2910–2921.

Zhang, L., & Wang, H. (2023). Chinese expert consensus on HIV pre-exposure prophylaxis (2023 Edition). *Chinese Journal of AIDS & STD, 29*(09), 954–961.

Zhou, T., Chen, Q., & Zhong, X. (2024). 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. *Annals of Epidemiology, 96*, 32–39.