AIDS Reviews

PERMANYER

Check for updates

REVIEW ARTICLE

Systematic review and meta-analysis of risk prediction models for HIV testing in key populations

Xin Xie[1,2#], Chaowen Zhang[3#], Shuyu Han[3], Yirong Shi[4], Jinyang Ming[3], Weimei Chen[4], Jingxian Hu[2], Bo Zhou[3], and Lili Zhang[2*]*

[1]School of Nursing, Capital Medical University, Beijing; [2]Department of Nursing, Beijing Youan Hospital Affiliated with Capital Medical University, Beijing; [3]Department of Nursing, School of Nursing, Peking University, Beijing; [4]Department of Infection and Immunology, the Third People's Hospital of Shenzhen, Shenzhen. China

#These authors contributed equally to this study and shared the first authorship.

Abstract

HIV testing is a critical tool for preventing HIV transmission, with early identification in key populations reducing onward spread. This study aims to evaluate risk prediction models, identify factors influencing HIV testing, and provide recommendations to enhance testing among key populations. Electronic databases were searched for peer-reviewed and gray literature published in English and Chinese from January 1, 1996, to November 14, 2025. Two reviewers independently assessed methodological quality and extracted data. The prediction model risk of bias assessment tool was used to evaluate bias and applicability. Of 3693 initially identified studies, seven met the inclusion criteria. Reported area under the curve (AUC) values ranged from 0.72 to 0.82, and the pooled AUC of validated models was 0.77 (95% confidence interval: 0.70-0.84), indicating moderate discriminative performance. However, most of the studies were assessed as having a high risk of bias, primarily due to insufficient reporting in the analysis domain, limiting the reliability of existing models for clinical or public health applications. Predictors of HIV testing were broadly grouped into sociodemographic, behavioral, knowledge-related, and structural factors, and predictors of HIV testing varied considerably across models, reflecting differences in study populations, behavioral characteristics, and contextual factors across settings. Overall, although existing models demonstrate moderate predictive ability, their methodological limitations and lack of external validation restrict their generalizability and practical utility. Therefore, reliable prediction models remain limited. Future research should develop high-quality models with larger sample sizes, robust designs, and multi-center external validation to support clinical application, improve practical relevance, and inform strategies to advance HIV testing. Beyond HIV testing prediction, this study highlights the need for broader prevention approaches, including sexuality education, risk awareness, and safe sexual behaviors, alongside targeted testing interventions.

Keywords: *HIV/AIDS. HIV testing. Risk prediction model. Key populations. Systematic review.*

Correspondence:
Lili Zhang
E-mail: zhanglili2020@mail.ccmu.edu.cn
Shuyu Han
E-mail: 2116393033@bjmu.edu.cn

Received: 17-03-2026
Accepted: 04-05-2026
DOI: 10.24875/AIDSRev.26000004

Available online: 09-06-2026
AIDS Rev. 2026;28(2):56-70
www.aidsreviews.com

Introduction

HIV infection continues to pose a significant global public health challenge. By the end of 2024, an estimated 40.8 million individuals were living with HIV worldwide, with approximately 91.4 million cumulative deaths attributed to AIDS-related causes.[1] Despite notable advancements in HIV prevention and treatment, new infections persist disproportionately among key populations.[2] It is estimated that around 70% of new HIV cases globally occur within key populations and their sexual partners.[3] Key populations are defined as individuals at heightened risk of HIV infection due to specific behaviors,[4] including gay men and other men who have sex with men sex workers (MSM), female sex workers (FSW), transgender women (TGW), people who inject drugs (PWID), prisoners, and other incarcerated individuals.[5] Importantly, the magnitude of HIV risk differs across these populations. The risk of HIV acquisition is approximately 13 times higher among MSM and 22 times higher among PWID compared with the general population. Among FSW, the global HIV prevalence is around 12%, reaching up to 30.7% in high-prevalence settings. TGW are also disproportionately affected, with a pooled HIV prevalence of 19.1%.[6,7] In addition, HIV prevalence in prison populations is estimated to be 2-10 times higher than that in the general population. The increased susceptibility of these populations arises from the interplay of behavioral risk factors and structural determinants, including high-risk sexual behaviors and drug-related practices, such as inconsistent condom use, multiple sexual partners, and drug use involving unsafe injection practices.[8-10] Failure to identify HIV infection early and interrupt transmission within these groups not only heightens onward transmission through high-risk sexual behaviors and needle sharing,[11] but also delays diagnosis and treatment, leading to disease progression, poorer clinical outcomes, and a higher likelihood of undiagnosed infections persisting in the community.[12] Consequently, early identification and targeted prevention efforts among key populations are essential for effective HIV prevention and control.

HIV testing is crucial for detecting HIV infection, facilitating the diagnosis, treatment, and monitoring of AIDS, and determining an individual's HIV status.[13] It is the cornerstone of HIV diagnosis, treatment initiation, and long-term care, serving as a fundamental entry point for the HIV care continuum.[14] In addition, HIV testing plays a vital role in achieving the Joint United Nations Programme on HIV/AIDS (UNAIDS) 95-95-95 targets, which aim to eliminate the AIDS epidemic as a public health threat.[15] Enhancing HIV testing uptake among key populations enables individuals to become aware of their HIV status, reduce high-risk behaviors, and commence antiretroviral therapy (ART) early, thereby reducing HIV transmission at both individual and population levels.[16] Despite the integration of key populations into national HIV testing services guidelines in many countries, evidence indicates that HIV testing coverage among these groups remains inadequate on a global and regional scale.[17] Evidence from a systematic review and meta-analysis further illustrates this gap,[18] reporting that the median proportion of individuals tested for HIV within the past 12 months was 54% among MSM, 61% among sex workers, 49% among PWID, and 45% among transgender individuals. These findings highlight substantial variability in testing uptake across key populations and underscore the need for more targeted and context-specific HIV testing strategies. Such strategies should be guided by country-specific epidemiological profiles, including the distribution and relative burden of HIV across different at-risk populations, as well as local patterns of transmission, to more effectively address unmet testing needs.[19] Nowadays, various HIV testing approaches, including provider-initiated testing, community-based testing, and HIV self-testing, have been introduced to enhance access to testing.[20] However, the number of HIV diagnoses falls short of global targets, which may reflect gaps in timely detection among key populations, including low perceived risk, limited awareness of HIV status, concerns about confidentiality, fear of stigma in healthcare settings, and structural barriers to accessing testing services.[21,22] Consequently, promoting early diagnosis and reducing transmission within key populations poses significant challenges in HIV prevention and control efforts.[23] It is imperative to implement effective strategies to expand HIV testing opportunities for these specific groups and investigate the factors influencing HIV testing among key populations to improve testing coverage and awareness within these communities.

Previous studies have identified various factors associated with HIV testing uptake among key populations,[24-26] including sociodemographic characteristics, sexual behaviors, substance use, risk perception, and access to healthcare services. However, analyses that focus solely on individual influencing factors may not adequately capture the complex and multi-factorial nature of HIV testing behaviors.[27] In contrast, risk prediction models integrate multiple predictors simultaneously to estimate the probability of a specific

outcome, thereby offering greater accuracy, practicality, and operational feasibility compared to single-factor analyses.[28,29] These models have been widely utilized in clinical and public health research to predict disease occurrence and other health-related outcomes.[30,31] With advancements in big data analytics and statistical learning methods, the approaches for developing prediction models have become increasingly diverse, leading to improved predictive performance.[32] Furthermore, many models can be translated into simple, user-friendly tools, such as scoring systems or digital applications, which are non-invasive, low-cost, and highly acceptable to users.[33] These tools do not require sophisticated equipment, pose no physical harm to individuals, and have the potential to enhance efficiency while reducing societal and healthcare costs.[34,35]

In recent years, risk prediction models have gained traction among key populations, primarily to evaluate the risk of HIV infection and the probability of undergoing HIV testing. For instance, one study created a prediction model to estimate the risk of HIV, HCV, and syphilis infection among key populations in two Canadian provinces, achieving an area under the curve (AUC) of 0.79.[36] Another study focused on developing a risk prediction model for HIV infection among MSM and identified four significant predictive variables.[37] However, no systematic review has thoroughly synthesized the evidence regarding these models for HIV testing within key populations. Consequently, this study aimed to conduct a systematic review to summarize the characteristics, performance, and methodological quality of existing risk prediction models for HIV testing that target key populations. Employing the Joanna Briggs Institute systematic evaluation methodology,[38] this review seeks to provide a comprehensive overview of the current state of research and to inform future model development and strategies to enhance HIV testing among key populations.

Methods

This study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines,[39] and the study protocol was registered in PROSPERO (registration number: CRD42024504653, https://www.crd.york.ac.uk/prospero/search).

Search strategy

Our search began with an exploration of studies on MEDLINE (PubMed) to develop our search strategy. After identifying MeSH terms and free-text keywords, we customized them for each database before conducting a comprehensive search on each. The snowballing method was used to screen references cited in the included articles when necessary. Studies from databases such as China National Knowledge Infrastructure, Wanfang Database, PubMed, Web of Science, and ProQuest were retrieved since the onset of the AIDS epidemic in 1996 until November 14, 2025. Publications were restricted to English or Chinese. Detailed retrieval strategies can be found in the supplementary data. The key items of our systematic review are as follows:

- P (population): this study primarily focused on key populations, including but not limited to drug users, LGBTQ+ individuals, and sexual partners of people living with HIV (PLWH).
- I (intervention): risk prediction models were developed and validated for HIV testing among key populations.
- C (comparator): there was no competing model.
- O (outcome): the outcomes are centered on HIV testing behaviors, irrespective of testing methods, including serological testing, hospital-based testing, and self-testing.
- T (timing): not limited.
- S (setting): not limited.

Inclusion and exclusion criteria

The inclusion criteria for the study were as follows: (1) studies focused on individuals aged 18 years and older, including but not limited to drug users, LGBTQ+ individuals, and sexual partners of PLWH; (2) risk prediction models in the studies were validated either internally or externally; (3) study outcomes that concentrated on HIV testing behaviors, and (4) the study types included quantitative studies, encompassing but not limited to cross-sectional and cohort studies.

The exclusion criteria were (1) unavailability of the full text; (2) duplicate or overlapping publications, and (3) articles published in languages other than English or Chinese.

Study selection and screening

We employed NoteExpress (NE) to select the literature. Initially, all retrieved literature was imported into NE for deduplication. Subsequently, two researchers (XX and WZ) independently and rigorously screened the literature based on the established inclusion and exclusion criteria. The authors first assessed the titles

and abstracts, followed by a comprehensive review of the full texts that met the criteria. Furthermore, the reference lists of all eligible studies were examined to identify any potentially relevant studies. In cases of disagreement regarding study selection, a discussion among the three authors (XX, WZ, and SH) was conducted to achieve consensus.

Data extraction

Two authors (XX and WZ) independently extracted the data, resolving any disagreements through discussion until a consensus was reached, or by consulting a third author (SH) when necessary. Following the CHARMS checklist,[40] the following data were extracted from each study: (1) demographic information, including the country of data collection, setting, data source, study design, and outcome definition; (2) methods for handling missing data, predictor selection, modelling, and model validation and presentation, and (3) prediction outcomes, including accuracy, sensitivity, specificity, and AUC.

Quality assessment

We utilized the available version of the prediction model risk of bias assessment tool (PROBAST)[41] to assess the applicability and potential bias risk of the included studies. The PROBAST checklist serves as a critical appraisal tool for studies involved in the development, validation, or updating of prediction models for individualized predictions. It consists of 20 signaling questions categorized into four domains: participants, predictors, outcomes, and analysis.

Data synthesis and statistical analysis

A meta-analysis of the AUC values from the models was performed using R software (version 4.4.0) along with the R packages “metaphor” and “meta.” The statistical heterogeneity of the pooled rate was evaluated using the I^2 statistic and a corresponding p-value, where a $p < 0.05$ indicated the presence of heterogeneity. The I^2 values were categorized to indicate low, moderate, and high heterogeneity, with thresholds set at 25%, 50%, and 75%, respectively.[42] The selection between fixed-effects and random-effects models depended on the level of heterogeneity observed in the analysis outcomes.

Results

Search and study selection

Figure 1 depicts the selection and inclusion process utilized in this review. The initial search produced 3,693 articles; following the removal of duplicates, 3,541 papers remained. Subsequently, 3,464 articles were excluded based on title and abstract screening, resulting in 77 articles for full-text evaluation. After assessing the full texts, 74 articles were excluded: 47 studies were eliminated for not establishing prediction models or focusing solely on risk factors, 17 studies did not match the review’s population criteria, 10 studies presented outcomes inconsistent with the review’s objectives, and 4 studies meeting the inclusion and exclusion criteria were obtained from citations. Ultimately, seven studies[43-49] were included in this review.

Study characteristics

Table 1 presents a summary of the characteristics of the studies incorporated in this systematic review, encompassing publications from 2012 to 2025 conducted in the United States,[44,45] Tanzania,[46] China,[49] Ethiopia,[47] and sub-Saharan Africa.[43,48] Among these studies, four were prospective,[43,44,46,49] while three were retrospective.[45,47,48]

The study populations exhibited significant diversity. Brown et al.[43] concentrated on the sexual partners of PLWH recruited from outpatient sexually transmitted infection clinics, whereas Pan et al.[44] focused on substance users using data from a previous clinical trial. Two studies specifically addressed missed or delayed HIV diagnosis: Weissman et al.[45] examined individuals with missed opportunities for earlier HIV diagnosis using electronic health records, and Xu et al.[49] explored individuals with late HIV diagnosis using a national surveillance database. The remaining three studies targeted broader high-risk populations. Chikusi et al.[46] recruited clients engaged in index testing and partner notification services. Alie et al.[47] analyzed adolescents and young adults aged 15-24 years using nationally representative demographic and health survey data, while Jaiteh et al.[48] investigated adults who had ever undergone HIV testing using repeated cross-sectional survey data. Across all the studies included, HIV testing or diagnosis was the primary outcome. Nevertheless, there was considerable variation in the event definitions and sample sizes, with

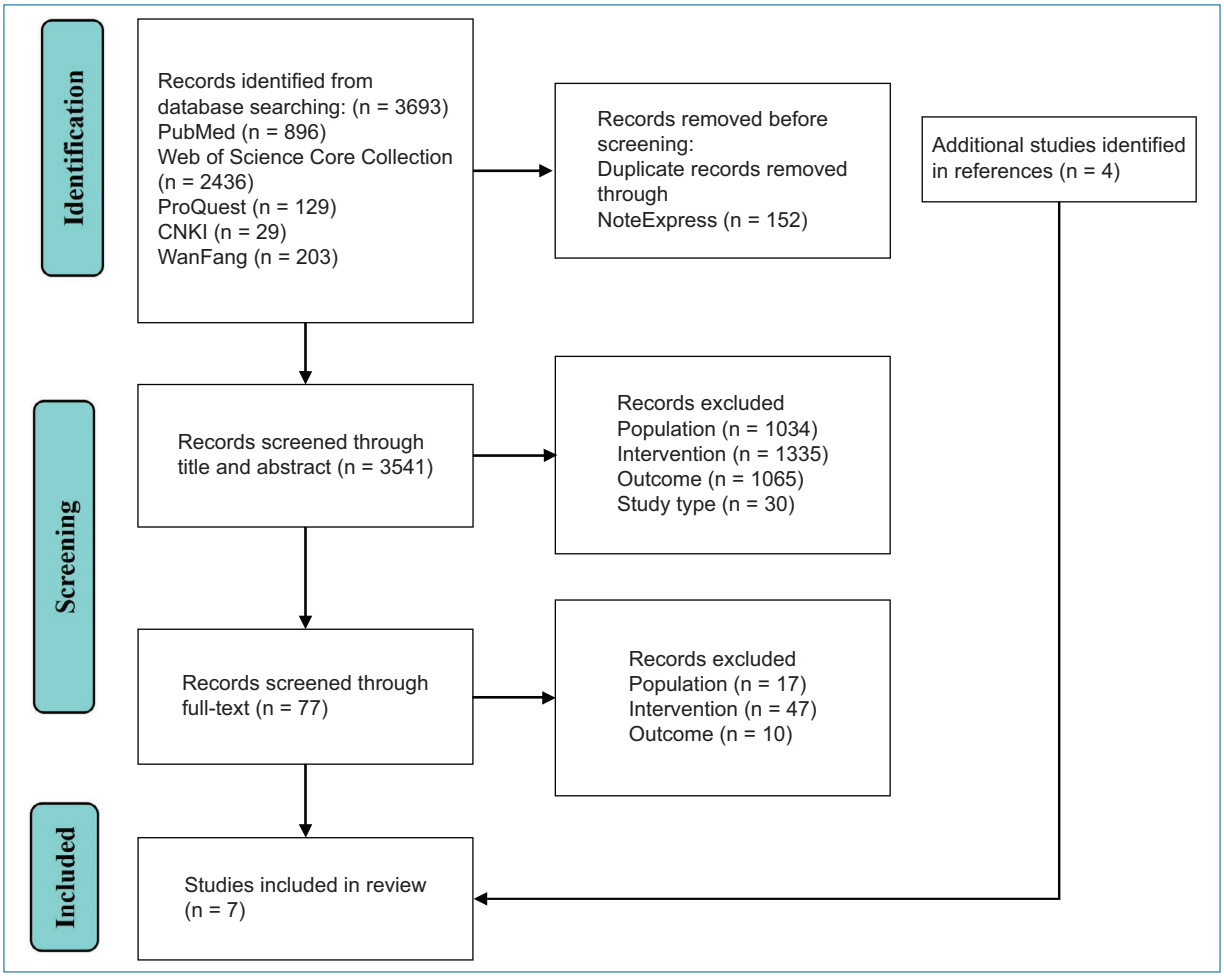

Figure 1. Preferred reporting items for systematic reviews and meta-analyses flowchart of literature search and selection.

event-to-sample ratios ranging from 4.4 to over 158.4/parameter, indicating substantial heterogeneity in study scale and statistical power.

Model characteristics

Table 2 summarizes the characteristics of the seven prediction models. One study utilized traditional logistic regression,[43] while the other six studies employed various machine learning techniques, such as random forest (RF), least absolute shrinkage and selection operator (LASSO)-based models, and other supervised learning methods.[44-49] The selection of candidate predictors varied among studies. Three studies defined candidate predictors primarily based on prior knowledge or existing literature,[44,45,48] while two studies included all available variables without specific pre-selection criteria.[46,47] The methods for selecting final predictors were inconsistently reported, including backward selection,[43] LASSO regularization,[45] univariable associations,[49] and bootstrap-based approaches,[44,47] with unclear procedures in one study.[46]

Across the seven included studies, no single predictor was consistently retained in all final models.[43-49] Sex and age were the most frequently included demographic variables, each appearing in four studies,[43,45,46,48] although their operationalization varied across models. Behavioral factors related to sexual activity and relationship characteristics were also commonly identified, particularly in studies focusing on individual- or community-level HIV testing.[43,44,47,48] In contrast, predictors reflecting HIV-related knowledge, media exposure, and healthcare access tended to emerge in models based on population surveys or programmatic data.[46-48] In addressing missing data, Pan et al.[44] utilized a non-parametric imputation

Table 1. Characteristics of the studies included in the systematic review

Author, Year	Study region	Study design	Participants	Data source	Main outcome	Events/Sample size (EPV or per parameter)
Brown et al., 2012[43]	Sub-Saharan Africa	Prospective cohort study	Sexual partner provided by PLWH within the last 3 months	The Kamuzu Central Hospital and the Bwaila Hospital outpatient sexually transmitted infection clinics	Sexual partner HIV testing and counseling	37/170 (4.6)
Pan et al., 2017[44]	America	Prospective cohort study	People ≥ 18 years or older and reported negative or unknown HIV status	National Institute on Drug Abuse Clinical Trials Network HIV testing and counseling study (CTN-0032)	HIV testing in substance users	763/1281 (4.4)
Xu et al., 2019[49]	China	Prospective cohort study	Late-diagnosed cases of HIV/AIDS	The Comprehensive AIDS Prevention and Control Information System of China from 2007 to 2017	The situation of late detection of HIV infection and its influencing factors	187/628 (17.0)
Weissman et al., 2021[45]	America	Retrospective study	Individuals at least 18 years of age who were diagnosed with HIV for the 1^{s} time between January 2008 and December 2016	Electronic health records data	Individuals with missed opportunities for earlier HIV diagnosis	1514/4725 (60.6)
Chikusi et al., 2022[46]	Tanzania	Prospective cohort study	The clients who participate in the Index Testing Services and Partner Notification Services	Different health centres and community sites from Arusha, Kilimanjaro, and Manyara	The number of HIV Index testing	Unknown
Alie and Negesse, 2024[47]	Ethiopia	Retrospective study	Individuals aged 15-24 years old	Publicly available data set of the demographic and health survey of 2016	Whether the adolescent had undergone HIV testing within the past 5 years	1553/4506 (50.1)
Jaiteh et al., 2025[48]	South Africa	Retrospective study	Male and female participants aged 18 years and older who provided a definitive response (Yes or No) to the outcome variable ("ever_hiv_test") during the survey	Five secondary datasets (South African national HIV prevalence, incidence, behaviour and communication survey [SABSSM] 2002, 2005, 2008, 2012, and 2017), which were originally conducted by the Human Science Research Council (HSRC)	Identify factors associated with HIV testing among South African adults aged 18 and older	2002: 1627/6228 (30.7) 2005: 4816/14285 (90.9) 2008: 6222/11833 (105.9) 2012: 16133/24263 (153.4) 2017: 26674/35071 (158.4)

method integrated into a RF algorithm, whereas Chikusi et al.[46] dealt with missing values by imputing the most frequent category. Alie and Negesse[47] employed multiple imputation for handling missing data, while Jaiteh et al.[48] utilized single imputation. The remaining three studies did not disclose either the proportion of missing data or the methods employed to address this issue.[43,45,49]

Model performance was reported inconsistently across studies. Discrimination was evaluated in most cases using the C-statistic or AUC, with values ranging from 0.72 to 0.82, which indicates a moderate level of discriminative ability. Internal validation was conducted in the majority of studies, primarily through cross-validation or bootstrap resampling. However,

Table 2. Characteristics of the models included in the systematic review

Author, Year	Number (%) and handling of missing data	Selection of candidate predictors	Selection of final predictors	Modelling method	Type of validation	Final predictors	Model performance	Model presentation
Brown et al., 2012[43]	n (%): unknown Method: no information	All available predictors	Backwards selection	Logistic regression	Int: bootstrap Ext: no information	Male partner gender, relationship duration < 6 months, relationship duration 6-24 months, and greater than primary education in the enrolled PLWH.	Cal: likelihood ratio tests Disc: C-Statistic/ AUC graph Ov: not evaluated	Summing the aOR to obtain a risk score
Pan et al., 2017[44]	n (%): unknown Method: nonparametric imputation method built in random forest	Based on prior knowledge	Bootstrap selection	Machine learning techniques	Int: cross-validation Ext: no information	HIV rapid testing on site, condomless sexual activity, etc.	Cal: not evaluated Disc: not evaluated Ov: not evaluated	No information
Xu et al., 2019[49]	n (%): unknown Method: no information	Based on univariable associations	Logistic regression	Machine learning techniques	Int: none (apparent performance) Ext: no information	Sample source, infection route, and age of diagnosis.	Cal: specify/0.839 Disc: C-Statistic/ AUC graph Ov: not evaluated	No information
Weissman et al., 2021[45]	n (%): unknown Method: no information	Based on prior knowledge	LASSO selection	Machine learning techniques	Int: cross-validation Ext: no information	Model 1: ED visit, older age, inpatient visit, alcohol use, male gender Model 2: older age, emergency medicine visit, internal medicine visit, male gender, alcohol use.	Cal: not evaluated Disc: C-Statistic/ AUC graph Ov: not evaluated	Formula of risk score obtained by Poisson LASSO coefficients of each factor
Chikusi et al., 2022[46]	n (%): unknown Method: the most occurring feature-filled, the missing values	All available predictors	Unclear	Machine learning techniques	Int: cross-validation Ext: no information	People with no knowledge, position of the client in society, marital status, age, and sex.	Cal: not evaluated Disc: not evaluated Ov: not evaluated	No information
Alie and Negesse, 2024[47]	n (%): unknown Method: multiple imputation	All available predictors	Bootstrap selection	Machine learning techniques	Int: cross-validation Ext: no information	Age, know place of HIV testing, age at first sex, recent sexual activity, exposure to family planning, number of sexual partner, media exposure, residence, educational level, awareness on AIDS, feature value, awareness on sexually transmitted infection, knowledge on contraceptive, knowledge on ovulatory cycle, literacy, wealth index, marital status, alcohol drinking, intimate partner violence, household members, number of children, feature value.	Cal: not evaluated Disc: C-Statistic/ AUC graph/0.809 Ov: 0.813	Integrating the machine learning algorithms with providing detailed individual-level data and incorporating this approach into community-based or facility-based testing programs
Jaiteh et al., 2025[48]	n (%): unknown Method: single imputation	Based on prior knowledge	Chi-square test	Machine learning techniques	Int: cross-validation Ext: unclear	Knowing a partner's HIV status, the place where to take an HIV test, an individual's level of education, receiving HIV education through the media (radio, TV, internet), sexual behaviors, being female, and being a younger adult.	Cal: not evaluated Disc: C-Statistic/ AUC graph Ov: not evaluated	Summing the OR to obtain a risk score

Int: internal validation; Ext: external validation; Cal: calibration measures; AUC: area under the curve, including Calibration plot, Calibration slope, Calibration-in-the-large (CITL), Hosmer Test, Other; Disc: discrimination measures, including C-statistic, D-statistic, AUC graph, log-rank test (if survival test), Risk group curves (if survival test), Other; Ov: overall measures, including R-squared, Brier score, Other.

none of the models underwent external validation. Calibration was largely overlooked, with only one study providing a calibration-related metric,[49] without adequate methodological detail. In terms of model presentation, only three studies converted model outputs into practical risk scores or formulas for potential clinical or public health applications.[43,45,48] The remaining studies offered limited insights into how the models could be implemented in real-world HIV testing programs.

Risk of bias and applicability assessment

The results of the risk of bias and applicability assessment are summarized in table 3. Overall, significant methodological limitations were identified across the included studies. Regarding the overall risk of bias, five studies were classified as high risk,[43-46] while two studies were deemed unclear.[47,48]

Risk of bias domains

In the participants' domain, three studies were assessed as high risk due to concerns regarding retrospective data sources or participant representativeness.[45,46,49] Two studies were rated as unclear due to insufficient information about participant recruitment or source populations.[43,47] The remaining two studies were evaluated as having a low risk of bias.[44,48] In the predictor's domain, two studies were classified as high risk, primarily due to non-transparent reporting or the potential influence of outcome information on predictors.[45,49] In addition, in several instances, predictors were selected based on univariable screening or prior knowledge without adequate justification, which increased the risk of biased estimation. Three studies were assessed as unclear due to inadequate reporting on the definition, processing, or handling of predictors in the presence of missing data.[43,44,46] The remaining two studies exhibited a low risk of bias.[47,48]

In the outcome domain, two studies were deemed to have a high risk of bias due to insufficient justification of outcome definitions or potential overlap with predictor information.[45,49] Three studies were classified as unclear because of limited reporting on outcome assessment procedures.[46-48] The remaining two studies were rated as having a low risk of bias.[43,44] The analysis domain emerged as the most problematic aspect across the included studies. Only one study was assessed as having a low risk of bias in this domain.[48] Four studies were rated as having an unclear risk of bias, primarily due to inadequate reporting on data complexity, the handling of missing data, and the processing of categorical and continuous variables.[45-47,49] Two studies were judged to have a high risk of bias, reflecting concerns related to predictor selection based on univariable screening, an insufficient number of outcome events relative to the number of predictors, and an increased risk of model over-fitting.[43,44]

Applicability assessment

Most of the studies exhibited relatively low concerns regarding applicability. In the participants' domain, all studies were rated as having low applicability concerns, indicating that the included populations were generally relevant to the review question. Similarly, concerns related to predictors were uniformly low across all studies. In the outcome applicability domain, three studies were rated as unclear due to insufficient reporting on outcome relevance or operationalization.[43,46,49] The remaining four studies were assessed as having low concerns, as the outcomes were well-aligned to predict HIV testing or diagnosis.[44,45,47,48]

Meta-analysis of the models included in the review

Among the seven articles ultimately included, four models reported AUC values and their 95% confidence interval (CI), and were suitable for meta-analysis. Utilizing a random-effects model, four validation datasets from three studies were synthesized to evaluate model discrimination. The pooled AUC was 0.77 (95% CI: 0.70-0.84), indicating acceptable discriminative performance of HIV testing risk prediction models across key populations. However, high heterogeneity between studies was observed (I^2 = 78.2%, τ^2 = 0.0015, Q = 13.76, p = 0.0032) (Fig. 2). Individual study AUC estimates varied from 0.72 to 0.82, with all CIs remaining above 0.70. The overall pooled effect was statistically significant (t = 35.83, p < 0.0001).

Discussion

This is the first systematic review focused on risk prediction models for HIV testing in key populations. The study aimed to identify factors influencing HIV testing within these populations and evaluate current modelling methodologies. This study lays the groundwork for research aimed at promoting HIV testing among key populations while also reducing the burden of early diagnosis and transmission of HIV.

Table 3. Risk of bias and applicability assessment results of the included studies

Author, Year	Risk of bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
Brown et al., 2012[43]	‡	‡	*	†	*	*	‡	†	‡
Pan et al., 2017[44]	*	‡	*	†	*	*	*	†	*
Xu et al., 2019[49]	†	†	†	‡	†	*	*	†	†
Weissman et al., 2021[45]	†	†	†	‡	*	*	*	†	*
Chikusi et al., 2022[46]	†	‡	‡	‡	*	*	‡	†	‡
Alie and Negesse, 2024[47]	‡	*	‡	‡	*	*	*	‡	*
Jaiteh et al., 2025[48]	*	*	‡	*	*	*	*	‡	*

*Indicates low ROB/low concern regarding applicability.
†Indicates high ROB/high concern regarding application.
‡Indicates unclear ROB/unclear concern regarding applicability.

Key populations are the primary drivers of new HIV infections.[14] Consequently, it is both essential and practical to prioritize and actively encourage these groups to undergo HIV testing. In contrast, the United Kingdom implemented an "opt-out" strategy for universal HIV testing across all populations in 2021.[50] Although this approach has effectively identified a greater number of undiagnosed infections and reduced HIV-related stigma, our findings indicate that further research is necessary to assess its effectiveness and long-term sustainability. This is particularly critical, as many regions, including countries such as China, remain low-prevalence settings where large population sizes present significant challenges to the implementation of comprehensive testing strategies.[51] Furthermore, universal testing may be challenging to operationalize in resource-limited settings, such as various areas of the Democratic Republic of the Congo, where the availability of HIV test kits is constrained for multiple reasons.[52]

The processes of study selection, data extraction, and quality assessment in this review also merit consideration. Following the literature screening, four models were identified that aligned with our study objectives. Although the number of available predictive models remains limited, many existing studies have focused primarily on identifying associated factors rather than developing validated prediction tools.[53-55] This highlights a clear gap in predictive models specifically tailored to key populations, particularly for HIV testing behaviors. Our evaluation indicated that the four models included in the study exhibited acceptable discriminative performance, with AUC values ranging from 0.72 to 0.82. However, most of the models were assessed as having a high risk of bias, primarily due to inadequate reporting of analytical methods, which highlights ongoing methodological challenges in model development.

The meta-analysis of four validation datasets produced a pooled AUC of 0.77, indicating moderate-to-good discrimination and supporting the potential utility of risk prediction models in identifying factors that may influence HIV testing behaviors.[56,57] Nonetheless, the wide CI (0.70-0.84) and significant heterogeneity suggest that model performance varies considerably across studies and settings, necessitating further model refinement before widespread implementation.[58] Moreover, AUC reflects discrimination only and does not capture calibration, clinical utility, or decision-analytic performance.[59] The lack of reported calibration metrics and impact analyses in several studies further limits conclusions regarding real-world applicability.

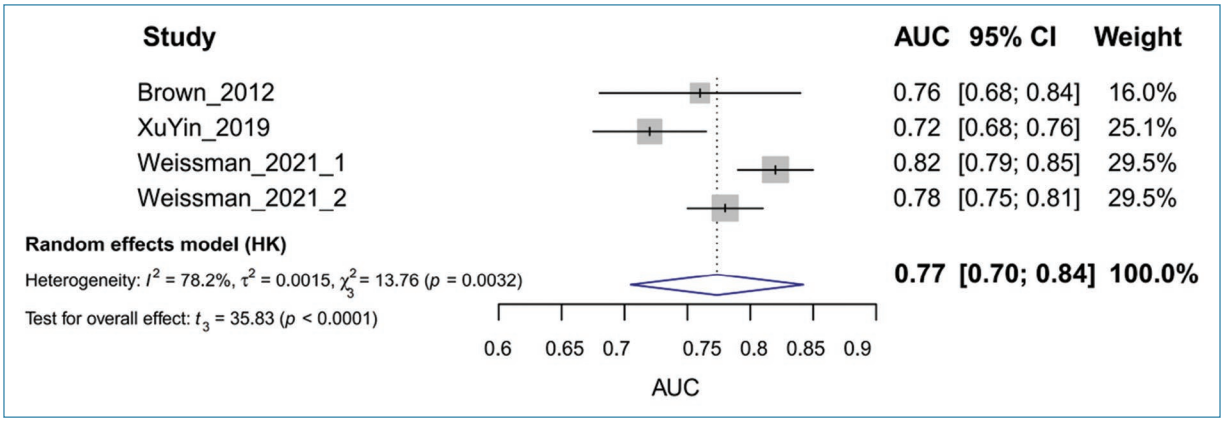

Figure 2. Forest plot of the random effects meta-analysis of pooled AUC estimates for three models.

The observed heterogeneity (I^2 = 78.2%) likely arises from variations in study populations, epidemiological contexts, and modelling strategies, including differences in key population definitions, HIV prevalence, predictor selection, and validation methods.[60-62] In addition, the similarity in weights and overlapping CIs between two validation datasets from the same study may further contribute to this heterogeneity.[45,63] Such factors limit the independence of pooled estimates and highlight the lack of external validation evidence.[64] Research indicates that harmonizing outcome definitions, predictor selection, and validation approaches can reduce heterogeneity and enhance comparability across studies.[65] Consequently, future research should prioritize the development of high-quality predictive models that focus on key populations and HIV testing behaviors. These initiatives must encompass larger sample sizes, robust study designs, transparent reporting of calibration measures, implementation-focused impact analyses, and multi-center external validation of existing HIV testing risk models across diverse key populations and settings.[66,67]

From the retained predictors in the models analyzed, researchers identified four broad categories of predictors present in multiple studies: Sociodemographic characteristics,[43,45,46,48,49] sexual behavior-related factors,[43,44,47,48] HIV-related knowledge, awareness, and information access,[46-48] and healthcare access or structural factors.[44,45,48] Sociodemographic factors were the most commonly reported predictors across studies; only one study excluded sociodemographic variables from its final HIV testing prediction model,[44] while the remaining studies retained at least one sociodemographic predictor.[43,45-49] This finding is somewhat inconsistent with existing evidence that demonstrates strong associations between HIV testing behavior and sociodemographic factors,[13] underscoring the need for further validation to clarify the role and relative contribution of these predictors across different contexts. In addition, although the included studies generally provided clear definitions and measurement approaches for their predictors, the validity and generalizability of the final predictors remain uncertain, because none of the seven studies conducted external validation, highlighting the necessity for future research to assess these predictors in independent datasets and diverse settings.[68,69] Meanwhile, the predictors in HIV testing prediction models varied significantly across studies, indicating a lack of universally shared predictors. This diversity suggests that HIV testing behavior is highly context-dependent, and predictors identified in one setting may not be directly applicable to others, particularly when differences exist in study populations, data sources, and testing methodologies.[70,71] Specifically, the included studies were primarily conducted in the United States[44,45] and African regions.[43,46-48] and can be broadly classified into resource-rich and resource-limited settings based on resource availability. Such regional differences may influence HIV testing behaviors through variations in cultural norms, socioeconomic conditions, and healthcare system capacity. For instance, resource-rich settings such as the United States generally have more developed healthcare infrastructure, greater availability of testing services, and fewer structural barriers to care than resource-limited settings. In this context, predictors identified in U.S.-based models were more likely to relate to healthcare utilization (e.g., emergency or inpatient visits), demographic characteristics (e.g., age and gender), and individual behavioral factors such as alcohol use

and sexual risk behaviors, suggesting a stronger link between testing and healthcare engagement.

In contrast, in resource-limited settings in parts of Africa, predictors tended to reflect a broader range of socioeconomic, informational, and structural factors, including knowledge and awareness of HIV and testing services, educational level, media exposure, and indicators of socioeconomic position (e.g., wealth index and residence). Relational factors, such as marital status, number of sexual partners, and awareness of a partner's HIV status, were also commonly identified, indicating that testing behaviors in these settings may be more influenced by access to information, social context, and structural inequalities. Cultural beliefs, stigma, and policy environments (e.g., routine or opt-out testing) may further shape these regional differences.[72,73] However, it should also be noted that this classification may not fully reflect within-country heterogeneity.

In addition to geographical variation, heterogeneity in study populations may further contribute to the observed variability in predictors. The included studies encompassed diverse participant groups, including sexual partners of PLWH,[43] adults with unknown HIV status (substance users),[44,48] individuals with delayed HIV diagnosis,[45,49] participants engaged in index testing and partner notification services,[46] and adolescents.[47] These populations differ substantially in risk exposure, prior testing experience, health awareness, and social and relational contexts, which may shape the relative importance and direction of predictors across models.

For example, individuals engaged in partner notification services may exhibit higher awareness and testing uptake due to structured interventions and provider-initiated contact;[74] accordingly, predictors in this group are more likely to reflect basic sociodemographic and structural characteristics, such as age, sex, marital status, and social position, rather than individual-initiated health-seeking behaviors. In contrast, sexual partners of PLWH may be more influenced by interpersonal dynamics and perceived risk within intimate relationships,[73] with predictors more likely to center on relationship-related factors, including partner gender, relationship duration, and educational level.

Adolescents may be more affected by developmental factors, health literacy, and prevailing social norms; therefore, their predictors are more likely to involve a broader combination of knowledge- and behavior-related factors, such as awareness of HIV and sexually transmitted infections, sexual behaviors, access to information (e.g., media exposure), and educational attainment. Meanwhile, individuals with unknown HIV status or those with delayed HIV diagnosis may be more influenced by informational gaps, risk perception, and barriers to early engagement with testing services;[75] accordingly, their predictors often reflect healthcare access and utilization (e.g., availability of on-site rapid testing and prior healthcare visits), as well as individual risk behaviors such as condomless sexual activity and alcohol use.

In addition, two included studies[45,49] further illustrate variability in predictive factors at the model level when examining studies focusing on similar clinical populations across different settings. Specifically, in studies focusing on individuals with delayed HIV diagnosis, the Chinese study identified predictors such as sample source, infection route, and age at diagnosis, whereas the English study more frequently emphasized healthcare utilization and behavioral factors (e.g., emergency department visits, inpatient visits, alcohol use, and male gender). Although based on a limited number of studies, this comparison suggests that even within similar populations, the relative importance of predictive factors may vary across healthcare systems and epidemiological contexts.[76]

As a result, predictors identified in one population may not be directly transferable to others. Taken together, these findings highlight the importance of accounting for both contextual and population-level heterogeneity when interpreting predictive models. Future research may benefit from more standardized definitions of study populations, stratified analyses within more homogeneous subgroups with appropriate adjustment for potential confounding factors, or the incorporation of contextual variables to improve the comparability, robustness, and generalizability of predictive models.

The included studies employed traditional regression methods[43] and machine-learning techniques for model development and predictor selection.[44,45] While both approaches are capable of developing prediction models, evidence suggests that machine-learning methods may offer superior accuracy compared with traditional regression.[77,78] A growing trend toward integrating machine learning into modelling has been observed, with its application expanding across diverse domains, including artificial intelligence-driven computer-aided drug design and development, disease prediction, and clinical decision support.[79-81] Relevant policies[82] further emphasize the importance of incorporating artificial intelligence (AI) and medical technology

on a global scale. Consequently, future research should place greater emphasis on exploring machine-learning approaches to enhance model performance.

This study establishes a foundation for the development of models in future research. The interventions discussed in the included studies also provide important insights for subsequent investigations. For example, one study suggested the implementation of targeted referral services,[43] while another recommended extending on-site counseling sessions and following up with individuals who declined services to comprehend their reluctance and offer tailored solutions.[45] These recommendations serve as valuable references for guiding future intervention research. Therefore, future research should leverage the model development experiences outlined in this review, incorporate established predictive factors and proposed intervention strategies, and translate these elements into enhanced HIV testing risk prediction models to more effectively promote HIV testing uptake among key populations.

Beyond HIV testing, a more comprehensive prevention framework is needed, incorporating sexuality education, affective awareness, and safe sexual practices. These factors play a critical role in shaping individuals' risk perception, sexual decision-making, and health-seeking behaviors. As highlighted by Soriano et al.,[83] sexuality is influenced by a complex interplay of biological, psychological, and sociocultural factors. During adolescence, individuals often experience evolving emotional and sexual perceptions, which may affect their understanding of identity, relationships, and sexual behaviors. In this context, providing accurate, age-appropriate information and fostering open communication with parents, educators, and healthcare professionals are essential to support healthy development. Such approaches may contribute to more informed decision-making and reduce engagement in high-risk sexual practices.[84]

Importantly, these gaps in sexuality and affectivity education are not only relevant to HIV transmission but also extend to a broader spectrum of sexually transmitted infections (STIs). As noted by Soriano et al.,[85] insufficient understanding of sexual health may increase vulnerability to multiple STIs. HIV and other STIs share common transmission routes and are often driven by similar behavioral factors, including inconsistent condom use, multiple sexual partnerships, and substance use during sexual activity.[8] In addition, contemporary social environments, including digital and online interactions, may further influence adolescents' sexual behaviors by facilitating partner seeking and increasing exposure to diverse sexual content, potentially shaping norms and risk perception.[86]

Therefore, prevention strategies should not address HIV in isolation but rather adopt an integrated sexual health approach. Strengthening comprehensive sexuality education, promoting healthy relationships, improving risk perception, and encouraging protective behaviors such as consistent condom use remain essential, alongside ensuring access to testing, counseling, and other supportive services.[87] Overall, while HIV testing remains a cornerstone of current prevention strategies, it should be complemented by education- and behavior-oriented interventions. Integrating these approaches may not only support earlier identification of infection but also address upstream determinants of sexual risk behaviors, thereby providing a more effective and sustainable pathway for reducing both HIV and other STIs at the population level.

Limitations

This study has several limitations. First, the inclusion of non-English literature may introduce potential language-related bias; however, it also allowed for the inclusion of additional context-specific evidence. Second, most included studies were conducted in the United States and African regions, which may introduce regional bias due to differences in healthcare systems and resource availability. Further research is needed in underrepresented regions, particularly in Asia and Europe. Third, the included studies covered diverse key populations with limited representation within each subgroup, which constrained the ability to conduct stratified analyses and may affect the generalizability of the findings. Finally, the relatively broad time span of the included studies may have influenced the identification and relevance of predictive factors, given that HIV testing practices, policies, and technologies have evolved. However, no substantially outdated studies were included, which partially mitigates this concern.

Conclusion

This systematic review identified seven studies developing risk prediction models for HIV testing in key populations. Overall, most models were assessed as having a high risk of bias according to PROBAST, indicating substantial methodological limitations in the current evidence base. In addition, the heterogeneity in study populations and settings suggests that predictive

factors may vary across different contexts, particularly between resource-rich and resource-limited regions.

These findings highlight the need for future research to develop more rigorous and generalizable prediction models. Efforts should focus on larger and more representative samples, improved study design, and robust external validation across diverse populations and settings. Strengthening methodological quality will be essential to enhance the reliability, clinical applicability, and potential impact of prediction models in supporting HIV testing strategies.

Acknowledgments

We are grateful for the support in conducting this systematic review offered by all the co-authors.

Authors' contributions

X. Xie: writing – original draft, formal analysis, data curation, conceptualization. C. Zhang: writing – original draft, formal analysis, data curation. S. Han: methodology, conceptualization, writing – review & editing. Y. Shi: methodology, data curation, funding acquisition. J. Ming: formal analysis, data curation. W. Chen: data curation. J. Hu: resources, investigation. B. Zhou: investigation. L. Zhang: writing – review & editing, supervision.

Funding

This study was supported by Capital's Funds for Health Improvement and Research (CFH) (grant number: 2024-2-2183). This study was funded by the Shenzhen Science and Technology Innovation Program. (JCYJ20230807143409019).

Conflicts of interest

None.

Ethical considerations

Protection of human subjects and animals. The authors declare that no experiments on humans or animals were performed for this research.

Confidentiality, informed consent, and ethical approval. This study does not involve personal patient data, medical records, or biological samples, and does not require ethical approval. SAGER guidelines do not apply.

Declaration on the use of artificial intelligence. The authors declare that no generative artificial intelligence was used in the writing or creation of the content of this manuscript.

Supplementary data

Supplementary data are available at DOI: 10.24875/AIDSRev.26000004. These data are provided by the corresponding author and published online for the benefit of the reader. The contents of supplementary data are the sole responsibility of the authors.

References

1. UNAIDS. UNAIDS; 2026. Available from: https://www.unaids.org/en/homepage [Last accessed on 2026 Feb 03].
2. Motaharinezhad M, Yousefi Z, Rostami S, Goli S, Keramat A. Reproductive health needs of HIV serodiscordant couples: a systematic review. Front Public Health. 2024;12:1348026.
3. Rocha GM, Cândido RC, De Carvalho NP, Carvalho EG, Costa AA, Machado IV, et al. Strategies to increase HIV testing among men who have sex with men and transgender women: an integrative review. BMC Infect Dis. 2023;23:240.
4. Larki M, Bahri N, Latifnejad Roudsari R. Life loaded with threat and vulnerability: a qualitative inquiry into the experiences of HIV negative married women in serodiscordant heterosexual relationships. BMC Womens Health. 2021;21:402.
5. Rucinski K, Banda LM, Olawore O, Akolo C, Zakaliya A, Chilongozi D, et al. HIV testing approaches to optimize prevention and treatment for key and priority populations in Malawi. Open Forum Infect Dis. 2022;9:ofac038.
6. WHO. Consolidated Guidelines on HIV Prevention, Diagnosis, Treatment and Care for Key Populations. Available from: https://www.who.int/publications/i/item/9789241511124 [Last accessed on 2026 Apr 17].
7. WHO. Consolidated Guidelines on HIV Prevention, Testing, Treatment, Service Delivery and Monitoring: Recommendations for a Public Health Approach; 2021. Available from: https://www.who.int/publications/i/item/9789240031593 [Last accessed on 2026 Apr 17].
8. Lee P, Docrat A. Prevalence and shared risk factors of HIV in three key populations in Vietnam: a systematic review and meta-analysis. Epidemiol Infect. 2023;151:e138.
9. Phuthomdee S, Teeraananchai S, Triamwichanon R, Benjarattanaporn P, Boettiger DC, Phanuphak N. HIV incidence and associated risk factors among key populations accessing routine testing through the national health coverage system in Thailand: a nationwide real-world study. Ann Epidemiol. 2025;110:178-86.
10. Kigombola A, Lyimo J, Mizinduko M, Mkembela D, Maziku E, Kafura W, et al. Low engagement of key populations in HIV health services in Tanzania: analysis of community, legal and policy factors. Pan Afr Med J. 2023;45:8.
11. Fraser H, Borquez A, Stone J, Abramovitz D, Brouwer KC, Goodman-Meza D, et al. Overlapping key populations and HIV transmission in Tijuana, Mexico: a modelling analysis of epidemic drivers. AIDS Behav. 2021;25:3814-27.
12. Kuchukhidze S, Walters MK, Panagiotoglou D, Boily MC, Diabaté S, Russell WA, et al. The contribution of intimate partner violence to vertical HIV transmission: a modelling analysis of 46 African countries. Lancet HIV. 2024;11:e542-51.
13. Jiao K, Wei R, Li H, Chow EP, Piqueiras E, Lewis T, et al. HIV testing frequency and associated factors among five key populations in ten cities of China: a cross-sectional study. BMC Infect Dis. 2022;22:195.
14. Solomon SS, Solomon S, McFall AM, Srikrishnan AK, Anand S, Verma V, et al. Integrated HIV testing, prevention, and treatment intervention for key populations in India: results from a cluster randomised trial. Lancet HIV. 2019;6:e283-96.
15. UNAIDS. 2021 UNAIDS Global AIDS Update - Confronting Inequalities - Lessons for Pandemic Responses from 40 Years of AIDS. UNAIDS; 2021. Available from: https://www.unaids.org/en/resources/documents/2021/2021-global-aids-update [Last accessed on 2024 Apr 22].
16. Sundararajan R, Ponticiello M, Nansera D, Jeremiah K, Muyindike W. Interventions to increase HIV testing uptake in global settings. Curr HIV/AIDS Rep. 2022;19:184-93.

17. Onovo A, Kalaiwo A, Agweye A, Emmanuel G, Keiser O. Diagnosis and case finding according to key partner risk populations of people living with HIV in Nigeria: a retrospective analysis of community-led index partner testing services. EClinicalMedicine. 2022;43:101265.
18. Stannah J, Soni N, Lam JK, Giguère K, Mitchell KM, Kronfli N, et al. Trends in HIV testing, the treatment cascade, and HIV incidence among men who have sex with men in Africa: a systematic review and meta-analysis. Lancet HIV. 2023;10:e528-42.
19. Khandu L, Hallett J, Crawford G, Leavy JE, Vujcich D. HIV self-testing and partner notification strategies for key populations in low- to upper-middle-income countries: a mixed-methods systematic review. PLoS One. 2025;20:e0338639.
20. Patel SN, Emerson BP, Pitasi MA, Tripp NE, Marcus R, Delaney KP, et al. HIV testing preferences and characteristics of those who have never tested for HIV in the United States. Sex Transm Dis. 2023;50:175-9.
21. Mbeda C, Ogendo A, Lando R, Schnabel D, Gust DA, Guo X, et al. Healthcare-related stigma among men who have sex with men and transgender women in Sub-Saharan Africa participating in HIV prevention trials network (HPTN) 075 study. AIDS Care. 2020;32:1052-60.
22. Maksut JL, Sanchez TH, Wiginton JM, Scheim AI, Logie CH, Zlotorzynska M, et al. Gender identity and sexual behavior stigmas, severe psychological distress, and suicidality in an online sample of transgender women in the United States. Ann Epidemiol. 2020;52: 15-22.
23. Maruyama H, Franks J, Laki D, Msumi O, Makyao N, Rwabiyago OE, et al. Bringing HIV services to key populations and their communities in Tanzania: from pilot to scale. J Int AIDS Soc. 2021;24:e25718.
24. Mizinduko M, Moen K, Pinkowski Tersbøl B, Likindikoki SL, Alexander Ishungisa M, Leyna GH, et al. HIV testing and associated factors among female sex workers in Tanzania: approaching the first 90% target? AIDS Care. 2023;35:850-8.
25. Anbessa O, Tesema Z, Amante E, Dufera T, Birhanu A. HIV testing acceptance and its associated factors among key and priority populations in Harar town, Harari region, Eastern Ethiopia. BMC Infect Dis. 2025;25:598.
26. Moameri H, Shahrbabaki PM, Tavakoli F, Saberi P, Mirzazadeh A, Goudarzi R, et al. Facilitators and barriers of HIV pre-exposure prophylaxis use among four key populations in Iran. BMC Health Serv Res. 2024;24:1433.
27. Jiang Y, Li J, Lu J, He L, Hu L, He J, et al. Predicting HIV self-testing intentions among Chinese college students: a dual-model analysis integrating health belief constructs and machine learning prioritization. Front Public Health. 2025;13:1596876.
28. Jung MS, Dlamini NS, Cui X, Cha K. Prevalence of HIV testing and associated factors among young adolescents in Eswatini: a secondary data analysis. BMC Pediatr. 2022;22:659.
29. Mapingure M, Mukandavire Z, Chingombe I, Cuadros D, Mutenherwa F, Mugurungi O, et al. Understanding HIV and associated risk factors among religious groups in Zimbabwe. BMC Public Health. 2021;21:375.
30. Zhang Y, Gu K, Du W, Xu A. Risk factors and prediction model for mortality in HIV/*Talaromyces marneffei* co-infection: a retrospective cohort study. Heliyon. 2024;10:e32560.
31. Ibeneme SC, Odoh E, Martins N, Ibeneme GC. Developing an HIV-specific falls risk prediction model with a novel clinical index: a systematic review and meta-analysis method. BMC Infect Dis. 2024;24:1402.
32. Albernas A, Patel MD, Cook RL, Vaddiparti K, Prosperi M, Liu Y. HIV risk score and prediction model in the united states: a scoping review. AIDS Behav. 2025;29:2388-407.
33. Xian Z, Song X, Wang Y, Yang T, Mao N. Construction and validation of a nomogram to predict 1-year mortality risk in patients with HIV/AIDS undergoing maintenance hemodialysis. Ren Fail. 2025;47: 2461665.
34. Coomes D, Green D, Barnabas R, Sharma M, Barr-DiChiara M, Jamil MS, et al. Cost-effectiveness of implementing HIV and HIV/syphilis dual testing among key populations in Viet Nam: a modelling analysis. BMJ Open. 2022;12:e056887.
35. Apers H, Nöstlinger C, Beckhoven DV, Deblonde J, Apers L, Verheyen K, et al. Identifying key elements to inform HIV-testing interventions for primary care in Belgium. Health Promot Int. 2019;35:301-11.
36. Vialard F, Zhang Q, Webster D, Materniak S, Dumont Blais A, Nair S, et al. Developing and validating a Bayesian clinical risk prediction model for three sexually transmitted infections in key populations from two Canadian provinces. Sex Transm Infect. 2025;101:467-73.
37. Dong Y, Liu S, Xia D, Xu C, Yu X, Chen H, et al. Prediction model for the risk of HIV infection among MSM in China: validation and stability. Int J Environ Res Public Health. 2022;19:1010.
38. 1.5 Publishing a JBI Systematic Review - JBI Manual for Evidence Synthesis - Confluence. Available from: https://jbi-global.atlassian.net/wiki/spaces/manual/pages/355598492/1.5+publishing+a+jbi+systematic+review [Last accessed on 2024 Oct [27].
39. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
40. Fernandez-Felix BM, López-Alcalde J, Roqué M, Muriel A, Zamora J. CHARMS and PROBAST at your fingertips: a template for data extraction and risk of bias assessment in systematic reviews of predictive models. BMC Med Res Methodol. 2023;23:44.
41. Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170:51-8.
42. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557-60.
43. Brown LB, Miller WC, Kamanga G, Kaufman JS, Pettifor A, Dominik RC, et al. Predicting partner HIV testing and counseling following a partner notification intervention. AIDS Behav. 2012;16:1148-55.
44. Pan Y, Liu H, Metsch LR, Feaster DJ. Factors associated with HIV testing among participants from substance use disorder treatment programs in the US: a machine learning approach. AIDS Behav. 2017;21:534-46.
45. Weissman S, Yang X, Zhang J, Chen S, Olatosi B, Li X. Using a machine learning approach to explore predictors of healthcare visits as missed opportunities for HIV diagnosis. AIDS. 2021;35:S7-18.
46. Chikusi H, Leo DJ, Kaijage DS. Machine learning model for prediction and visualization of HIV index testing in Northern Tanzania. Int J Adv Comput Sci Applic. 2022;13:391-9.
47. Alie MS, Negesse Y. Machine learning prediction of adolescent HIV testing services in Ethiopia. Front Public Health. 2024;12:1341279.
48. Jaiteh M, Phalane E, Shiferaw YA, Jallow H, Phaswana-Mafuya RN. The application of machine learning algorithms to predict HIV testing using evidence from the 2002-2017 South African adult population-based surveys: an HIV testing predictive model. Trop Med Infect Dis. 2025;10:167.
49. Xu Y, Luo Y, Lu H, Tang J. Influencing factors of late diagnosis for newly identified HIV/AIDS cases in Jiangyin City and prediction model. Occup Health. 2019;35:1086-9.
50. The Lancet Hiv. Opt-out HIV testing in the UK. Lancet HIV. 2023;10:e351.
51. Wu Z, Chen J, Scott SR, McGoogan JM. History of the HIV epidemic in China. Curr HIV/AIDS Rep. 2019;16:458-66.
52. Migaud P, Naidenow J, Ghaeni L, Hosmann K, Kaghoma J, Kowomba J, et al. Indicator-condition-guided HIV testing in rural DRC. Infection. 2024;52:1171-4.
53. Heckman TG, Sikkema KJ, Kelly JA, Fuqua RW, Mercer MB, Hoffmann RG, et al. Predictors of condom use and human immunodeficiency virus test seeking among women living in inner-city public housing developments. Sex Transm Dis. 1996;23:357-65.
54. Johnson BA, McKenney J, Ricca AV, Rosenberg ES, Liu C, Sharma A, et al. Risk factors associated with repeated HIV testing among internet-using men who have sex with men. AIDS Educ Prev. 2016;28:511-23.
55. Wahome EW, Graham SM, Thiong'o AN, Mohamed K, Oduor T, Gichuru E, et al. Risk factors for loss to follow-up among at-risk HIV negative men who have sex with men participating in a research cohort with access to pre-exposure prophylaxis in coastal Kenya. J Int AIDS Soc. 2020;23 Suppl 6:e25593.
56. Buyukatak E, Anil D. An investigation of data mining classification methods in classifying students according to 2018 PISA reading scores. Int J Assess Tools Educ. 2022;9:867-82.
57. Çorbacıoğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: a guide to interpreting the area under the curve value. Turk J Emerg Med. 2023;23:195-8.
58. Van Leeuwen FD, Steyerberg EW, Van Klaveren D, Wessler B, Kent DM, Van Zwet EW. Instability of the AUROC of clinical prediction models. Stat Med. 2025;44:e70011.
59. White N, Parsons R, Collins G, Barnett A. Evidence of questionable research practices in clinical prediction models. BMC Med. 2023;21:339.
60. Elayan H, Sperrin M, Martin GP, Peek N, Braunschweig F, Faxén J, et al. Correcting for case-mix shift when developing clinical prediction models. BMC Med Res Methodol. 2025;25:186.
61. Yordanov TR, Lopes RR, Ravelli AC, Vis M, Houterman S, Marquering H, et al. An integrated approach to geographic validation helped scrutinize prediction model performance and its variability. J Clin Epidemiol. 2023;157:13-21.
62. De Jong VM, Moons KG, Eijkemans MJ, Riley RD, Debray TP. Developing more generalizable prediction models from pooled studies and large clustered data sets. Stat Med. 2021;40:3533-59.
63. Nikoloulopoulos AK. Joint meta-analysis of two diagnostic tests accounting for within and between studies dependence. Stat Methods Med Res. 2024;33:1800-17.
64. Bu Z, Sun Y, Shi Z, Hu L, Ju C, Liu Y, et al. Prognostic prediction models for ulcerative colitis: systematic review and meta-analysis. J Med Internet Res. 2025;27:e71944.
65. Lusa L, Kappenberg F, Collins GS, Schmid M, Sauerbrei W, Rahnenführer J, et al. Evaluation of changes in prediction modelling in biomedicine using systematic reviews. BMC Med Res Methodol. 2025;25:167.
66. Archer L, Snell KI, Ensor J, Hudda MT, Collins GS, Riley RD. Minimum sample size for external validation of a clinical prediction model with a continuous outcome. Stat Med. 2021;40:133-46.
67. Routier L, Touati S, Ghostine-Ramadan G, Wallois F, Querne L, Bourel-Ponchel E. External validation of a model for predicting outcomes in preterm newborns. JAMA Netw Open. 2025;8:e2523029.

68. Mansmann U, Ön BI. The validation of prediction models deserves more recognition. BMC Med. 2025;23:166.
69. Marconi L, Cabitza F. Show and tell: a critical review on robustness and uncertainty for a more responsible medical AI. Int J Med Inform. 2025;202:105970.
70. Leber W, Farooq HZ, Panovska-Griffiths J, Larvin H, Baggaley RF, Divall P, et al. Risk prediction models for targeted testing of HIV, hepatitis B and hepatitis C: a systematic review and meta-analysis. BMC Infect Dis. 2025;25:1648.
71. Luo Q, Luo Y, Cui T, Li T. Performance of HIV infection prediction models in men who have sex with men: a systematic review and meta-analysis. Arch Sex Behav. 2023;52:2011-23.
72. Sabo KG, Seifu BL, Kase BF, Asebe HA, Asmare ZA, Asgedom YS, et al. Factors influencing HIV testing uptake in Sub-Saharan Africa: a comprehensive multi-level analysis using demographic and health survey data (2015-2022). BMC Infect Dis. 2024;24:821.
73. Mannoh I, Amundsen D, Turpin G, Lyons CE, Viswasam N, Hahn E, et al. A systematic review of HIV testing implementation strategies in Sub-Saharan African Countries. AIDS Behav. 2022;26:1660-71.
74. Jamal A, Srinivasan M, Kim G, Nevins AB, Vohra S. Factors associated with HIV testing within the national health interview survey (2006-2018). J Racial Ethn Health Disparities. 2024;11:2638-48.
75. Ahmed N, Ong JJ, McGee K, D'Elbée M, Johnson C, Cambiano V, et al. Costs of HIV testing services in sub-Saharan Africa: a systematic literature review. BMC Infect Dis. 2024;22:980.
76. Croxford S, Burns F, Copas A, Yin Z, Delpech V. Trends and predictors of linkage to HIV outpatient care following diagnosis in the era of expanded testing in England, Wales and Northern Ireland: results of a national cohort study. HIV Med. 2021;22:491-501.
77. Song X, Liu X, Liu F, Wang C. Comparison of machine learning and logistic regression models in predicting acute kidney injury: a systematic review and meta-analysis. Int J Med Inform. 2021;151:104484.
78. Zhao X, Lu Y, Li S, Guo F, Xue H, Jiang L, et al. Predicting renal function recovery and short-term reversibility among acute kidney injury patients in the ICU: comparison of machine learning methods and conventional regression. Ren Fail. 2022;44:1326-37.
79. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Mol Divers. 2021;25:1315-60.
80. Lee YW, Choi JW, Shin EH. Machine learning model for predicting malaria using clinical information. Comput Biol Med. 2021;129:104151.
81. Peiffer-Smadja N, Rawson TM, Ahmad R, Buchard A, Georgiou P, Lescure FX, et al. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. Clin Microbiol Infect. 2020;26:584-95.
82. WHO. WHO Issues First Global Report on Artificial Intelligence (AI) in Health and Six Guiding Principles for its Design and use; 2021. Available from: https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use [Last accessed on 2024 Mar 24].
83. Soriano V, Vélez O, Sánchez-Méndez JI, González-Rico N, González L. Research on differences in sexual orientation and identity. AIDS Rev. 2025;27:163-70.
84. Powell M, Krentz HB, Eagles ME, Gill MJ. Missed opportunities within healthcare for an earlier diagnosis of HIV. Int J STD AIDS. 2020;31:1169-77.
85. Soriano V, Blasco-Fontecilla H, Gallego L, Fernández-Montero JV, De Mendoza C, Barreiro P. Rebound in sexually transmitted infections after the COVID-19 pandemic. AIDS Rev. 2023;26:127-35.
86. Gyane CO, Gmayinaan VU, Osei E. Association between social media use and adolescents' sexual behaviours: a cross-sectional study among high school students in Drobo, Ghana. BMC Public Health. 2025;25:537.
87. Bordogna AL, Coyle AC, Nallamothu R, Manko AL, Yen RW. Comprehensive sexuality education to reduce pregnancy and STIs in adolescents in the United States: a systematic review and meta-analysis. Am J Sex Educ. 2022;18:39-83.