Development of a Predictive Model for Sexually Transmitted Infections in Individuals Using Pre-Exposure Prophylaxis for HIV in Spain

Principal Investigator: Anaïs Corma Gómez
Version 3
Date: 11/02/2025

Background

Pre-exposure prophylaxis (PrEP) is an effective and safe strategy for HIV prevention in men who have sex with men (MSM), transgender women, and cisgender heterosexual individuals (1-4). The daily combination of two antiretroviral drugs, tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC), in a single oral tablet, has been authorized as a preventive treatment for HIV at both the European and national levels. It was included in the pharmaceutical benefits covered by the Spanish National Health System in 2019. Although the implementation of PrEP programs in Spain has been slower compared to neighboring countries and was further delayed by the SARS-CoV-2 pandemic, all Autonomous Communities (CCAA) currently provide this service. According to data from the Information System of PrEP Programs in Spain, as of September 2021, there were 9,092 individuals on PrEP, primarily men, especially MSM. While initially targeted at MSM, transgender individuals, and cisgender women engaged in sex work, the eligibility criteria were expanded at the end of 2021 to allow any individual over 16 years old in a situation of vulnerability to HIV to request PrEP. The estimated number of users taking PrEP as of April 2023 was 18,075, according to aggregated reports from the CCAA (6).

PrEP programs are not solely based on pharmacological measures but also incorporate other interventions such as condom promotion, sexual education, assisted counseling, substance use assessment, risk reduction, early detection and treatment of sexually transmitted infections (STIs), and vaccination status updates (5). All these activities are conducted quarterly in medical, nursing, and pharmacy consultations within Specialized Care Units for Infectious Diseases and Hospital Pharmacy Departments. The eligibility criteria include having more than 10 different sexual partners, engaging in condomless anal sex, using drugs associated with condomless sex (chemsex), having used post-exposure prophylaxis (PEP) on multiple occasions, and having been diagnosed with at least one bacterial STI in the past year. Therefore, individuals at high risk of acquiring HIV infection also share a similarly high risk of other STIs.

Despite its effectiveness in reducing HIV incidence, the high prevalence of sexually transmitted infections (STIs) among individuals using PrEP presents additional challenges in terms of public health and clinical management. While HIV infections in this context remain rare, STI prevalence remains significant, underscoring the need to address this public health concern comprehensively. In fact, STIs are the most frequently diagnosed clinical event during the follow-up of PrEP users. Several studies conducted in international cohorts have estimated a very high incidence of STIs in this population (7-10). However, the impact of PrEP on STI incidence trends remains debated. While some studies suggest that initiation into the program leads to a decrease in condom use, thereby increasing the risk of acquiring an STI (11), others have found no such increase (9,10,12). Additionally, it is crucial to recognize that the PrEP-using population is

heterogeneous, meaning that not all individuals share the same risk of acquiring an STI. In fact, evidence suggests that approximately 75% of STIs occur in just 25% of the PrEP population (12).

In any case, STI screening is a fundamental pillar of the PrEP program and should be conducted at every clinical visit, even in asymptomatic patients, to diagnose, treat, and eliminate STIs, thereby interrupting the chain of transmission. However, there is widespread debate regarding the optimal screening frequency, as testing for certain STIs is costly, severe complications in men are rare, and frequent antibiotic use may contribute to antimicrobial resistance. On one hand, a Dutch modeling study suggested that testing for chlamydia and gonorrhea every three months, compared to every six months, would not be cost-effective according to current health economic criteria (13). On the other hand, two recent studies have demonstrated that biannual screenings could delay the diagnosis and treatment of a substantial proportion of asymptomatic STIs, potentially leading to increased transmission (14,15). A risk-based approach to determining STI screening frequency could help resolve these challenges.

Since the approval and subsequent implementation of PrEP in Spain at the end of 2019, the demand for these services has steadily increased. In fact, access to PrEP is estimated to be significantly lower than the existing demand. The growing need for services, along with the additional burden associated with STI monitoring, is beginning to overwhelm the resources of PrEP-providing units. The requirement for quarterly STI screenings, combined with the rising number and diversity of PrEP users, poses significant logistical and financial challenges for healthcare systems. For these reasons, the interaction between this strategy and STI incidence raises fundamental questions. In our setting, large-scale data on STI incidence among PrEP users, its distribution across different population profiles, and its trends over time remain unknown. Currently, all PrEP-providing units offer the same type of follow-up for all patients. In this context, there is an urgent need to develop innovative approaches to address and prevent STIs more efficiently within the PrEP population. Artificial intelligence (AI) can provide robust tools for the development of predictive models that integrate a wide range of risk factors, individual characteristics, and their dynamic changes. This would enable more accurate risk stratification and the personalization of preventive and follow-up interventions. A personalized approach, in which each individual is stratified according to their risk of experiencing the studied events, could aid in clinical decision-making. Moreover, this approach would not only facilitate the early identification of individuals at higher risk but also allow for the effective design of a tailored care and follow-up model for these patients. This would optimize resources, reduce the burden on both individuals and the healthcare system, and enhance efficiency while maximizing health outcomes.

Optimizing follow-up not only improves efficiency and reduces costs but also enhances the quality of care by focusing on personalization and active prevention. Ultimately, this project aims to make a substantial contribution to public health by improving the management of STIs in a specific and dynamic population.

References

1. Grant RM et al. NEJM.2010; 363:2587-99.

2. Grant RM et al. Lancet Infectious Diseases. 2014; 14: 820-9.

3. Baeten JM et al. NEJM. 2012; 369: 399-410.

4. Drak D et al. AIDS. 2021; 35: 2319-26.

5. GeSIDA. Recomendaciones sobre la Profilaxis Pre-Exposición para la Prevención de la Infección por VIH en España. Marzo 2023. Disponible en [http://.gesida-seimc.org/category/guias-clinicas/otras-guias-vigentes/].

6. División de Control de VIH, ITS, Hepatitis virales y Tuberculosis. Sistema de información de programas de Profilaxis Pre-exposición al VIH en España (SIPrEP). Informe de resultados noviembre 2019-mayo 2023. Ministerio de Sanidad, julio 2023. Disponible en [https://siprep.isciii.es/documentos].

7. Hoornenborg E et al. Lancet HIV. 2019; 6: e447-55.

8. Traeger MW et al. JAMA. 2019; 321 (14): 1380-90.

9. Traeger MW et al. Lancet Infectious Diseases. 2022; 22:1231-41.

10. Ayerdi O et al. EIMC. 2024 [Epub ahead of print] DOI number:10.1016/j.eimc.2024.01.003.

11. Traeger MW et al. CID. 2018; 67: 676-86.

12. McManus H et al. JAMA Network Open. 2020; 3 (12):e2030806.

13. Van Wifferen et al. Sex Transm Infect. 2021; 97: 607-12.

14. Tang EC et al. AIDS. 2020; 34: 1181-6.

15. Jongen VW et al. Sex Transm Infect. 2023; 99: 149-155.

Hypothesis

It is possible to accurately define the risk of acquiring a sexually transmitted infection (STI) in individuals using PrEP who are included in the follow-up program.

Objectives

➢ **Primary Objective:**

- Develop and validate a predictive model for STI acquisition in individuals using PrEP within the national provision program for this strategy.

➢ **Secondary Objectives:**

- Estimate the incidence of STIs in the cohort, both globally and by type of infection.
- Assess the evolution of STI diagnoses within the program and model their long-term trend, both globally and by specific pathogen.
- Evaluate the proportion of asymptomatic STIs diagnosed within the screening program.
- Study behavioral changes in the socio-sexual domain of individuals initiating PrEP.
- Identify potential candidates for the “Doxy-PEP” strategy (doxycycline use as post-exposure prophylaxis for chlamydia and syphilis).

Methods

- **Design and Patients:** This is a multicenter, ambispective cohort study involving 24 Spanish hospitals. All individuals using PrEP and enrolled in the follow-up program will be included (see PrEP criteria).
- **Follow-up:** The entry point into the program will be considered the baseline visit. From that point onward, patients will be followed quarterly according to national PrEP follow-up protocols in Spain. At each visit, the following assessments will be conducted:

  -Risk evaluation

  -Clinical and laboratory assessment

  -Sample collection for STI diagnosis, including: serum, pharyngeal swab, urethral exudate or urine sample, rectal swab

If patients present specific STI-related symptoms, targeted sample collection will be performed for diagnostic purposes. The presence of symptoms suggestive of an STI will be defined as the development of any of the following: mucocutaneous ulcerative or other lesions, dysuria, urethral discharge, proctalgia, pathological stool products, rectal tenesmus.

- **Variables**

➢ Outcome Variables:

Primary Outcome: Development of an STI (Yes/No). The following STIs will be considered: Syphilis, Neisseria gonorrhoeae (NG) infection, Chlamydia trachomatis (CT) infection, Lymphogranuloma venereum (LGV), Mycoplasma genitalium (MG) infection, Genital herpes, Hepatitis A, Hepatitis B, Hepatitis C, HIV infection, Mpox

Secondary Outcome: Number of STIs diagnosed

➢ Explanatory Variables:

Sociodemographic variables: Age, gender identity, sex, sexual orientation.

Variables related to sexual practices:

Receptive and/or insertive sex

Frequency of condom use

Number of sexual partners

Engagement in chemsex and route of drug use

Sex work and frequency of visiting sex exchange venues

Variables related to STIs:

Clinical presentation

Site of infection

Vaccination status for hepatitis A (HAV), hepatitis B (HBV), human papillomavirus (HPV), and smallpox

Variables related to PrEP use:

Previous use of PrEP and/or post-exposure prophylaxis (PEP)

PrEP regimen (continuous/on-demand)

Adherence to PrEP

- **Data Collection**

Clinical, epidemiological, analytical, and microbiological data will be recorded in a Case Report Form (CRF) specifically designed for this study. For data collection and management, the RedCap® platform has been selected due to its security, flexibility, and ability to structure databases in a standardized manner, facilitating both data monitoring and traceability. Data entry will be performed locally by professionals external to the research team, who provide this service to study participants. A semi-annual remote monitoring of all data recorded in the CRF will be conducted, along with on-site monitoring in specific cases, as determined by the data monitoring team.

- **Laboratory Determinations**

These determinations will be performed as part of routine clinical practice in the Microbiology Unit Laboratory of each participating center. Diagnosis of Chlamydia trachomatis (CT) and Neisseria gonorrhoeae (NG) will be conducted using nucleic acid detection techniques in the following sample types:

-Urine

-Endocervical swab

-Urethral swab

-Rectal swab

-Pharyngeal swab

Screening for HIV, syphilis, and viral hepatitis will be performed using serological techniques. Syphilis reinfection will be defined as a four-fold or greater increase in RPR titers. For individuals with positive anti-HCV antibodies, HCV RNA PCR testing will be performed to rule out active reinfection.

Statistical Analysis

Once recruitment is completed, the study population will be stratified into three subsets, ensuring that the distribution of patients with respect to the outcome variable remains similar across all groups. Two-thirds of the population will be used for developing predictive models, while the remaining third will serve as the validation group for model assessment.

- Incidence Calculation:
    - The quarterly incidence will be calculated by dividing the number of STI cases by the total number of individuals during that period.
    - Incidence rates will be expressed per 100 person-years of follow-up, with exact 95% confidence intervals (CIs).
    - Participants will contribute to person-time calculations until their last STI screening or until administrative censoring of follow-up, whichever occurs first.
    - Individuals who do not receive a PrEP prescription within three months after their previous prescription will be considered to have discontinued PrEP, and their follow-up will be censored three months after their last prescription.
    - If a participant resumes PrEP at a later stage, they will be re-entered into the analysis, with their person-time calculation restarting from their subsequent STI screening.
- Descriptive Analysis:
    - Categorical variables will be expressed as absolute and relative frequencies.
    - Continuous variables will be reported as medians (Q1–Q3).
    - Comparison of proportions: Chi-square test or Fisher’s exact test.

    - Comparison of continuous variables: Student's *t*-test or Mann-Whitney *U* test.

- Multivariate Analysis:
    - To identify factors associated with STI acquisition, both bivariate and multivariate logistic regression analyses will be performed.
    - Variables showing an association with $p < 0.050$ in the bivariate analysis will be included in the multivariate model.
    - The significance level will be set at $p < 0.05$, and results will be expressed as odds ratios (ORs) with 95% confidence intervals.
- Machine Learning Models for Predictive Analysis:
    - Various machine learning models will be trained to predict future STI infections in new patients.
    - While interpretable models (e.g., decision trees, rule-based models) will be prioritized, more complex models (e.g., ensemble methods and neural networks) will also be explored if they demonstrate significantly superior performance (Bonaccorso, G. et al., 2018).
    - Hyperparameter tuning will be performed using cross-validation on the training set.
    - The best-performing models will then be evaluated on the validation set to assess their ability to generalize to new patients not included in the training phase.
- Model Evaluation and Comparison:
    - Different performance metrics will be used, accounting for class imbalance and the differential costs associated with prediction errors (Raeder, T. et al., 2012).
    - Model comparison will follow a Bayesian approach, as suggested by Benavoli et al. (2017), or alternative statistical comparisons (Garcia, S. et al., 2008).
    - A Region of Practical Equivalence (ROPE) will be defined through a consensus between theoretical considerations and clinical expertise.
    - Among the best models falling within the ROPE, the most interpretable model will be selected.
- Software:

- Statistical and model-fitting analyses will be performed using:
    - SPSS version 26.0 (IBM Corporation, Somers, NY, USA)
    - STATA 16.0 (StataCorp LP, College Station, TX, USA)
    - Python (statsmodels, scikit-learn, and other relevant libraries) (latest stab

- **Ethical Considerations**

This research study will be conducted in accordance with the ethical principles outlined in the Declaration of Helsinki by the World Medical Association (WMA), as well as the following legal and regulatory frameworks: Law 14/2007 on Biomedical Research; Law 41/2002, the fundamental law regulating patient autonomy, rights, and obligations regarding clinical information and documentation; Organic Law 3/2018 on Personal Data Protection and Digital Rights; Good Clinical Practice (GCP) guidelines; Regulatory authority requirements for the verification of original documents and study audits/inspections. All collected data will undergo an encryption process to ensure patient anonymity. Additionally, approval for this study has been approved by the Research Ethics Committee (REC) of the Southern Health Management Area of Seville (coordinating center) and the respective ethics committees of all participating centers.