# Aim 1: Describe the types of PrEP persistence trajectories among male PrEP users in the U.S. through cluster analysis.

## 2.1 Introduction

Men having sex with men (MSM) are the population most affected by HIV in the United States (U.S.). In 2018, MSM accounted for around 69% (n=2,6198) of incident HIV diagnoses in the US [1], despites only comprising about 2% of the U.S. population. Therefore, it is crucial to improve the effectiveness of HIV prevention strategies targeted toward MSM in order to reduce the overall burden of new HIV infections in the US, which has otherwise not been meaningfully reduced since 2014 [2].

Since the approval of TDF-FTC (emtricitabine e/tenofovir disoproxil fumarate) by the Food Drug Administration in 2012, pre-exposure prophylaxis (PrEP) has been recommended for MSM at risk of HIV acquisition[3] due to its high efficacy in randomized controlled trials[4, 5]. Despite high efficacy in controlled settings, the effectiveness of PrEP is entirely dependent on user persistence over time. Based on findings from the open-label extension study of the iPrEX trial (a phase III PrEP trial among MSM), an adherence of at least 4 doses per week (or 57% of proportion of days covered in PrEP) is needed to achieve an optimal level of protection[6].

The lack of persistence in PrEP care over time posts a substantial challenge to its success as a HIV preventive measure among MSM. Data from PrEP clinics in three different cities indicated that 72% and 57% MSM users remained in PrEP care 3 months and 6 months after initiation, respectively[7, 8]. In addition, a multi-site review of clinical records suggests that only about 30% of PrEP-initiating MSM were retained at 12 months[9]. A follow-up survey of an open-label PrEP study at San Francisco and Miami found that only 40% of the respondents reported continual use of PrEP after the completion of the original study, even though 71% indicated very high interest in continuing with PrEP use after the study completion[10].

Although poor persistence in PrEP care within the first year of PrEP initiation has been well documented in the real world among MSM, few observational studies have analyzed persistence in PrEP care over a more extensive follow-up timeline. In addition, there have been various working definitions of PrEP persistence used in existing studies that analyze PrEP use over time (e.g., continued filling of PrEP prescription without a gap of >1 month at 3,6 and 12 months after initiation[11], having at least ¾ of 1 and 2 years with >50% of days in a month covered with PrEP supply[12], utilizing PrEP care in the last 6 month[13]). In many of these definitions, a person’s persistence in PrEP was determined by whether they met an arbitrary PrEP use threshold within a pre-specified, often broad, time intervals (e.g., 3 months, 12 months). It is often unclear how the implications of these definitions (for example, a person is “PrEP-persistent” at month 6 since initiation) are connected to the time-varying risk of suboptimal sero-protection, the understanding of which is vital to PrEP care case management. Furthermore, the broad time interval, in which PrEP persistence is determined, could increase the risk of ecological bias when it is used for inferring one’s PrEP use experience at more granular time points. As the result, there is a need to understand PrEP persistence by analyzing the longitudinal trajectory of PrEP use at a finer scale while considering the linkage between the dynamical patterns of PrEP use and their implication for the PrEP user’s sero-protection.

Furthermore, no existing studies to our knowledge have generated typologies of longitudinal PrEP persistence trajectories among the MSM populations in the U.S. From the perspective of PrEP care providers, it is important to leverage existing data to describe a typology of the longitudinal PrEP persistence trajectories because heterogeneity in PrEP use patterns among MSM may require tailored PrEP persistence support strategies. More specifically, such an understanding may inform clinicians about patterns of PrEP cessations and re-initiation, thereby aiding PrEP persistence interventions to target these key timings. In addition, an appropriately characterized typology of PrEP persistence patterns may facilitate patient segmentation, which in turn allows more efficient allocation of PrEP persistence intervention resources.

Thus, the overarching goal of this analysis is to describe the common patterns of longitudinal PrEP persistence using data-driven methodologies to identify latent clusters of PrEP persistence. We hypothesize that there is substantial heterogeneity in the longitudinal PrEP use behaviors such that we may qualitatively differentiate the mined patterns of PrEP persistence. Furthermore, we aim to identify key individual-level and neighborhood-level characteristics associated with identified PrEP persistence patterns.

Coy and Siegler found that age, health insurance characteristics and pharmacy type are associated with greater PrEP persistence (i.e., having >50% of a month period with PrEP supply for over 18 months) in the first two years since initiation among adult PrEP users in the U.S. Mental health conditions, particularly depression and attention deficit hyperactivity disorder (ADHD) have been identified as a risk factor for ART non-adherence in the existing literature[14]. A PrEP open label study (iPrEP OLE) found a modest association between depression severity and protective level of PrEP adherence in the MSM cohort[15]. Given the high prevalence of depression in MSM and the selective nature of prior clinical trial research[16, 17], gaining further understanding on how depression is associated with PrEP persistence in a general population study is critical. Similarly, past research on ADHD suggested that individuals with ADHD have greater tendency of medication non-adherence (e.g., diabetes patients[18], HIV patients [19]). No previous studies to our knowledge have explored the association between ADHD and PrEP persistence among MSM. The disparity in PrEP use by race/ethnicity[20] and neighborhood-level factors (e.g., socioeconomic economic status[21], PrEP service density[22]) and PrEP use have also been documented in existing literature.

Findings regarding the presence or absence of heterogeneity in patient characteristics by PrEP persistence pattern class will facilitate the hypothesis-generating process for future research that aims to understand mechanisms by which individual-level and neighborhood-level characteristics are associated with certain PrEP use patterns. The product of this analysis will enable development of adaptive PrEP persistence interventions by providing specific recommendations on the identification of groups at high risk for PrEP non-persistence (according to the PrEP persistence pattern) as the intervention target.

## 2.2 Methods

### 2.2.1 Data source & study population

The target population in this study is the adult (aged 18 – 65 years) MSM PrEP users in the US. To obtain a representative dataset for this population, we will extract about 6,000 simple random samples of all adult male users who had at least two years of longitudinal PrEP prescription fill records by 1/1/2020 (i.e., initiated on PrEP in 1/1/2015 ~ 1/1/2018) from a national commercial pharmacy. The inclusion criteria for a PrEP prescription fill record are:

1. Anti-retroviral monotherapy that consists of only TDF-FTC. This excludes prescriptions to treat HIV that are required to include other antiretroviral medications
2. To be identified as a prescription record associated with PrEP initiation, have at least a 60-day supply of TDF-FTC; less supply duration (e.g. 30 days) indicates HIV post-exposure prophylaxis.

In addition to prescription data, we will extract individual-level demographic, financial, and geographic information from the commercial pharmacy’s prescription database, in order to assess their association with longitudinal PrEP persistence patterns (Table 1). In addition, we will create numbers of ZIP-3-level characteristics for each study subject, using the geographic information (ZIP code) from the PrEP users, for the association analysis (Table 1). Variables in Table 1 are selected based on data availability and observed associations with PrEP use and HIV treatment adherence [14-23].

**Table 1.** Individual-level and neighborhood-level characteristics selected for the association analysis with PrEP use trajectory patterns.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable Name | Definition/Note | Variable Type | Level (if applicable) |
| **Individual-level characteristics. Source: Commercial pharmacy database** | | | |
| Age at PrEP initiation (yr) |  | Continuous |  |
| Year of PrEP initiation | May provide insights regarding the secular trend of PrEP use behaviors | Categorical | 2015, 2016, 2017, 2018 |
| Primary payer type | Source of payment used most frequently in the 2-year tracking period. If two sources of payments were received, the commercial or government (including VA/Tricare) insurance coverage will be prioritized over other copay assistance. | Categorical | 1. Commercial 2. Government (Medicaid, Medicare) 3. Tricare/Veteran administration 4. Others |
| Average copay per month (USD) | Average out-of-pocket payment per month over the duration PrEP medication was filled | Continuous |  |
| Pharmacy type |  | Categorical | 1. Community-based specialty pharmacy 2. Traditional retail pharmacy |
| PrEP provider type |  | Categorical | 1. Medical doctors 2. Physician assistants/nursing practitioner |
| Depression medication prescription | Precise medication inclusion criteria will be based on prior studies from the commercial pharmacy. | Categorical | Yes/No |
| ADHD medication prescription | Precise medication inclusion criteria will be based on prior studies from the commercial pharmacy. | Categorical | Yes/No |
| **ZIP-3-level characteristics. Source: US Census** | | | |
| Concentrations of black populations | Unit: per 100,000 people | Continuous |  |
| Concentration of Latinx/Hispanic population | Continuous |  |
| Concentrations of residents under poverty threshold | Continuous |  |
| Concentrations of residents uninsured | Continuous |  |
| Density of PrEP provider[24] | Continuous |  |

The advantage of using a national chain pharmacy’s administrative dataset is that the database captures detailed medication fill records that reflect PrEP use behaviors in the real world for an extensive user population across the US. (It had pharmacy locations in 49 states and 20 of the most populous cities as of 2019.). Furthermore, past studies have validated pharmacy refill data as an appropriate approach to evaluate medication adherence and persistence[25, 26]. On the other hand, an inherent limitation of the pharmacy administrative dataset is that sexual orientation is not collected by the pharmacy. As a result, we will not be able to further filter for MSM among the extracted male PrEP users. This may undermine the transportability of our analysis’ findings. However, such a limitation should not cause serious concerns for the relevance of our analysis, because it is estimated that majority of the PrEP-indicated population in the U.S. are MSM (71%)[27] and less than 1% of the heterosexual active adults reported PrEP use[28].

### 2.2.2 Data pre-processing

Proportions of days covered in medication (PDC) is a common metric for medication adherence in an observed interval [29]. We will repeatedly compute modified, biweekly PDC (*bwPDC*) in PrEP over time, starting from each subject’s initial PrEP initiation date. Using weekly moving strides, we will create a ~6000×104 time series matrix (Note: 52 weeks/year × 2 years = 104 weeks):

**Eq. 1:** where = index number for the two-week windows; = the number of days with PrEP intake.

**Table 2. Empty data matrix for proportion of days covered**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Subject ID | (Day 1 – 14) | (Day 8 – 21) | (Day 15 – 28) | … | (Day 722 – 735) |
| 1 |  |  |  | … |  |
| 2 |  |  |  | … |  |
| 3 |  |  |  | … |  |

The challenge of formulating bwPDC time series such that they reflect the realistic adherence experiences is that the prescription data cannot reflect actual PrEP dosing schedules. The American Men’s Internet Survey (AMIS) data and findings from a large online survey of MSM suggest that over 70% of the PrEP users follow daily dosing[30]. Therefore, for the base analysis, we will assume that each PrEP user is following daily dosing from the start date of a prescription fill interval, and distribute the medication supply associated with each of their prescription fill accordingly across their follow-up timeline. This would allow us to obtain the number of days with PrEP intake in a two-week window used in the calculation of bwPDCs (Eq. 1). Because we are ultimately interested in evaluating PrEP persistence by the presence or absence of sufficient sero-protection, the bwPDC time series will then be converted to the time series of PrEP sero-protection using the minimum effective adherence threshold (i.e., 4 pills per week) as a binary cutoff (1= bwPDC ≥ 57%; 0= bwPDC <57%).

### 2.2.3 Group-base trajectory modeling (GBTM)

GBTM is an application of finite mixture modeling that has been widely used for identifying latent distinct clusters of individuals following similar trajectories (i.e., repeatedly measured phenomenon that progress over time)[31]. Under the finite mixture modeling framework, GBTM assumes that the observed trajectories arise from the realization of a finite mixture of pre-specified distributions. A mixture distribution is a probability distribution derived from an assembly of other random variables, in which a data value is realized from a component random variable that is randomly selected from that assembly of random variables. The distribution of trajectories (i.e., ), under the aforementioned framework, may be described by the following formula:

**Eq. 2:**

Where ) denotes the time series of measurements on individual *i* over the follow-up period; denotes the membership probability in group *j* (the finite mixture is composed of J groups or component random variables) for individual *i*; denotes the likelihood of conditioned on membership in group *j* and the corresponding distribution parameter vector .

To reduce model complexity, GBTM (as most finite mixture models that handle longitudinal data) assumes conditional independence between measurements at different time points (for example, and ) given their group membership [31], such that the conditional likelihood function of may be written as ). Trajectories will be modeled as a polynomial function of time. GBTM can also be used to model the effect of hypothesized risk factors on group membership (i.e., , where is risk factor vector for individual *i*) under the assumption that data trajectories are independent of risk factors given group membership.

In this analysis, we will model the conditional distribution of PrEP persistence data given group membership under Bernoulli distribution. Logistic regression model will be used to model the presence of sufficient PrEP sero-protection in a two-week window:

**Eq. 3:**

Where denotes whether a PrEP user *i* is considered sufficiently protected by PrEP over the next two weeks at week *t* since the initiation of PrEP given he belongs to group *j*; denotes the number of weeks since the PrEP initiation date and denotes the order of the group-specific polynomial used to model potentially non-linear trajectories (we will explore 2nd and 3rd order of polynomial in the time variable)

### 2.2.4 Evaluation of PrEP trajectory cluster

A major challenge with any data mining analysis, including GBTM, is that there is no external data on ground-truth label that can be used for validating the detected cluster solution. As the result, we will consider the following three aspects when evaluating the quality of the discovered PrEP trajectory clusters (or groups) solutions: 1) internal validity; 2) cluster stability; 3) cluster interpretability.

In an ideal world, we would want to evaluate all possible mined cluster structures given specified hyperparameters, which include the number of PrEP use trajectory clusters to mine (K) and the assignment of group-specific polynomial order in the time variable (e.g., if K=3, polynomial order for group 1,2 and 3 could be 2,2, and 3 respectively). However, the number of models evaluated would have increased substantially, which may not be feasible given limited computational power. Furthermore, the qualitative interpretability of the mined PrEP use trajectory cluster would likely decrease as K increases.

To limit the number of evaluated models and maintaining cluster interpretability, we will only assess K=1~9. We will first use the 3rd polynomial order in the time variable for all K groups (polynomial parsimony will be assessed at a later stage). We will evaluate internal validity, cluster stability, and cluster interpretability sequentially. As we move through each evaluation phase, we will narrow down candidate models. The reason we evaluate cluster internal validity before cluster stability is that the method we will use for evaluating cluster stability (i.e., consensus clustering) is more computationally expensive than that for evaluating internal validity. Therefore, it is more suitable for assessment at a later phase where we will have fewer number of models to be evaluated.

#### 2.2.4.1 Phase I: internal validity

A cluster solution with good internal validity should maximize model fit while balancing model complexity (i.e., the greater the number of clusters to mine, the greater model complexity), and have a clear assignment of cluster membership. To assess model fit, we will compute and compare the Bayesian information criterion (BIC). Lower BIC reflects greater data likelihood penalized for model complexity. To assess the clarity of group assignment, we will compute the average posterior probabilities of group membership for each discovered group by averaging the posterior probabilities of group membership among subjects assigned to the same group (according to the maximum probability of group membership). Higher average posterior probabilities of group membership indicate clearer partition. In addition, clear group assignment may also be indicated by 1) a large ratio comparing the odds of model-predicted group assignment against the odds of random group assignment (known as “Odds of Correct Classification” in the group-based trajectory modeling textbook by Nagin 2005[32]), and 2) close correspondence between the estimated probability of a group membership and the proportion of study subject assigned to that group[33].

By the end of this evaluation phase, we will identify four K values to pass onto the next evaluation phase (cluster stability). To ensure the selected K-group models reflect greater internal validity than others, we will sort the candidate models by BIC in ascending order, and select the top four models that also have average posterior probabilities ≥ 0.70 [33-35] and “Odds of Correct Classification” ≥ 5 [33, 36, 37] for all discovered groups. To ensure cluster representativity, all cluster sizes for a selected model should at least account for 5% of the study population[38].

To assess the strength of evidence against the hypothesis of homogenous PrEP use trajectories, we will approximate the Bayes factors (using the formula[39]:) comparing each selected K-group model against the single-group model. The obtained Bayes factor may be interpreted as the ratio of marginal likelihood of a selected model to that of the single-group (i.e., homogenous trajectory) model.

#### 2.2.4.2 Phase II: cluster stability

After selecting the top four values of K based on the aforementioned internal validity criteria, we will proceed to the second phase of evaluation where we evaluate the selected K values by assessing their resulting cluster stability against sampling variability with consensus clustering validation [40]. Such an assessment will allow us to identify cluster solutions generated by overfitting models that describe random noises, rather than the true cluster structure. For each selected K value, we will resample, without replacement, a subset (80%) of the original dataset, stratified by individual-level characteristics (i.e., age, year of PrEP initiation, health insurance type) for 100 times and apply the k-group GBTM to each sample dataset (i.e., 100 clustering runs for each selected K value). The aforementioned subsampling scheme is preferred over bootstrapping, because bootstrapping may artificially inflate the cluster compactness through replicating identical data points.

Upon running a K-group GBTM on all the sample datasets, we will measure the agreement among the resulting cluster solutions over the sampled datasets. To achieve this, we will compute a N×N consensus matrix, originally introduced by Monti et al[40], that describes the pairwise probabilities of being assigned to the same cluster conditioned on being sampled in the same subset (Eq 4).

**Eq 4.**

Where N is the number of data points of the original dataset; H is the number of sample datasets; is the connectivity matrix specific to a clustering run on sample datasets (i.e., 1 if data point *i* and *j* are assigned to the same group; 0 otherwise); is the indicator matrix specific to a clustering run (i.e., 1 if data point *i* and *j* are both sampled in the sample dataset; 0 otherwise).

Because GBTM is a fuzzy clustering method (i.e., the models would output probabilities of group membership for each person), we propose to adapt the calculation of consensus probabilities (i.e., ) to the following formula:

Where H denotes the number of sample datasets that contains both individual i and j; K denotes the pre-specified number of groups for GBTM to mine. The equation above effectively describes the probability that two individuals belong to the same group, averaged across the sample datasets that contain both individuals.

Each consensus matrix will be used to plot the histogram of the empirical cumulative distribution (CDF) for the consensus probabilities (i.e., ). A stable K-group solution would result in clear bimodal distribution of the consensus probabilities and a CDF shape that approaches a step function (Figure 1[40]). We may also calculate the proportion of ambiguous clustering (PAC), the proportion of consensus probabilities that lie within a pre-specified interval indicative of ambiguous pairwise connectivity (e.g., 0.1 ~ 0.9). PAC is a consensus measure developed by Yasin et al to infer optimal K in a more reliable manner[41]. A lower PAC value, corresponding to flatter middle segment of CDF (more step-function-like), suggests better K for clustering.

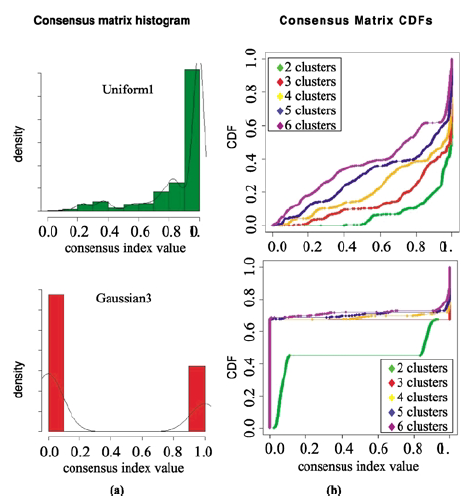
In addition, for a given K, we will also evaluate the stability of each resulting cluster (or group) by computing the average pairwise consensus probabilities (i.e., m(k)) among individuals belonging to the same group (Eq. 5). An ideal K-group solution should result in a cluster structure that have high and even m(k) across all clusters, suggesting that all discovered clusters are equally robust against data perturbation.

**Eq 5.**

Where denotes the number of individuals assigned to a group k; denotes the set of indices of individuals assigned to group k.

We will use the aforementioned consensus indicators along with internal validity measures from the first phase of cluster evaluation to guide us selecting two top performing K values. Then, for each selected K values, we will assess whether more parsimonious polynomial functions (2nd or 1st order) may be used to describe group-specific trajectories by comparing the Bayes factors between the full and reduced models.

**Figure 1.** Consensus for various number of groups (K) on the two simulated datasets (figure extracted from Monti et al 2003[40]). *Uniform1* was generated using uniform distribution (mimicking situations where there is no true cluster structure), while *Gaussian3* was generated using the union of three gaussian distributions (mimicking situations where there is a true cluster structure; =3 in this example).



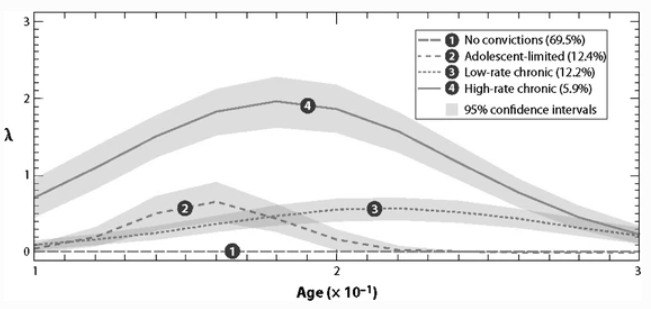
#### 2.2.4.3 Phase III: cluster interpretability

We will evaluate the cluster interpretability of the two top performing K-group solutions selected from the phase II cluster evaluation. From a data mining perspective, the ultimate goal of cluster analysis is to find a grouping or label structure that can provide meaningful, clinically relevant insights, given the domain understanding, on the differences between detected clusters. To facilitate this qualitative evaluation of the discovered PrEP persistent patterns, we will visualize the predicted trajectories of PrEP sero-protection over time by cluster membership (Figure 2). A group-specific predicted trajectory may be interpreted as the time-varying probabilities of suboptimal PrEP sero-protection in the two-week window, along the follow-up time of a typical PrEP user belonging to a trajectory group.

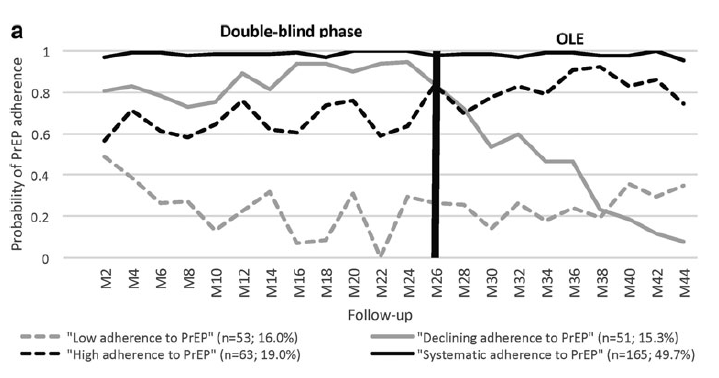
For each group’s predicted trajectory, I will identify the phases of PrEP cessation in which the probabilities of PrEP sero-protection consistently hover around or below 30% (the cutoff is informed by examining the predicted trajectories in Sagaon-Teyssier et al 2018[38](Figure 2b)), and describe their timings and durations. Events of possible re-initiation after cessation will also be identified. Describing the timings of probable PrEP cessations and re-initiation will allow us to better characterize each PrEP user cluster. We will select the K-group cluster solution with the most interpretable and distinguishable trajectories as the final cluster or pattern structure for the two-year PrEP persistence trajectories among male PrEP users in the U.S in the Pre-COVID era.

**Figure 2**. sample visualizations of model-predicted trajectories by cluster membership.

1. Trajectories of conviction counts among samples in London by detected and labeled clusters, using zero-inflated Poisson regression (Nagin et al 2010[42])



1. Trajectories of event-based PrEP adherence (i.e., at least one pill taken within 24 hours before and after sex) among MSM in the ANRS IPERGAY trial by detected clusters, using logistic model (Sagaon-Teyssier et al 2018[38])



### 2.2.5 Association analyses

We will obtain and compare the descriptive statistics of the selected individual-level and neighborhood-level characteristics (Table 1) by the final cluster membership. This will allow us to get an idea of the presence or absence of heterogeneity in these characteristics between PrEP users of different PrEP persistence patterns. We will also assess the bivariate association between the final PrEP use trajectory cluster membership and each of these characteristics. Because GBTM models fuzzy cluster membership (i.e., each data point has a probability of belonging to each cluster, as opposed to deterministically belong to only one cluster), we will account for the uncertainty of cluster membership in the bivariate association analyses by specifying a multinomial logistic function in the GBTM (Eq. 6).

**Eq. 6:**

Where *i* denotes subject index; *j* denotes the index for the mined clusters; denotes regression constant; denotes the coefficient of a characteristic variable; denotes the value of a characteristic variable for individual *i*; and are set to zero for identifiability[39].

Failure to account for uncertainty in cluster assignment may result in biased inference regarding the association between the mined PrEP use trajectory patterns and the selected characteristics [39, 43]. Through the model specified in Eq. 6, we will estimate the characteristics’ coefficients and the resulting multinomial odds ratios (with 95% confidence interval) as the associational measure. Through examining multinomial odds ratios, we will identify the key correlates of PrEP persistence pattern, which may further inform the development of future research that aims to explore the causal relationship between certain characteristics and PrEP persistence patterns.

### 2.2.6 Sensitivity analyses

We will conduct sensitivity analyses to assess the impact of uncertainty around the following factors on the detected cluster structure of PrEP persistence trajectories:

#### 2.2.6.1 The PrEP dosing schedules

We will evaluate the impact of non-daily dosing schedules on the mined typology from the group-based trajectory modeling (GBTM). First, we will define new terms:

* **PrEP assessment interval:** the time interval bounded by two consecutive PrEP prescription fill dates, or the end date of the two-year tracking period
* **Adherent PrEP consumption period:** the duration between the start date of a PrEP assessment interval and whichever comes earlier between the following two dates: 1) the end date of the PrEP assessment interval or 2) the end date of a 5-day extension period after the end of medications’ days supply. Given that most PrEP prescription fills contain 30-day supply, the 5-day relaxation corresponds to a dosing frequency of at least 6 doses per week, which fits the working definition of daily dosing schedule referenced in a nationwide survey study on PrEP use[30]. The adherent PrEP consumption period effectively represents the approximal duration a PREP prescription fill would last if a PrEP user mostly maintain daily dosing frequency.

We will assign 5% and 10% of PrEP users to follow nondaily dosing regimen. Non-daily PrEP users will be randomly selected from the subset of PrEP users who have at least one PrEP assessment intervals in which the interval lengths are greater than the length of the adherent PrEP consumption periods. We will assume that all PrEP users complete their prescribed supply before the next PrEP assessment interval for to preserve analysis parsimony. Through the aforementioned selection process, we will have two subpopulations of MSM:

1. **Daily PrEP users:** all of their PrEP assessment intervals would follow daily dosing regimen. For each prescription fill, the dates of PrEP intake (according to the medication days supply) will be randomly selected without replacement from the associated adherent PrEP consumption period.
2. **Non-daily PrEP users**: they would follow the daily dosing regimen for the PrEP assessment intervals in which the interval lengths were within the adherent PrEP consumption periods. However, they would use the non-daily PrEP regimen for the PrEP assessment intervals in which the interval lengths were greater than those of the adherent PrEP consumption periods. There are two types of non-daily PrEP users: event-based users and regular non-daily users. Each will account for 50% of the non-daily user population[30].
   * Event-based users: they will follow the on-demand 2-1-1 dosing regimen (i.e., two doses within 24 hours before sex and two single doses for the two days after sex), which has been shown to provide sufficient protection against HIV transmission among MSM[44-46]. For each prescription fill, we will assume that the event-based users will have sufficient sero-protection under on-demand dosing regimen, starting from the prescription fill date, until the date when the medication supply runs out under the average frequency of 1.5 sex per week, estimated among MSM in the U.S [47]. For example, an event-based user has a 30-day PrEP prescription fill at Day 1. He will be considered protected in Day 1~ Day 35 (i.e., the duration of sero-protection period=). We assume that all event-based users adhere to the 2-1-1 dosing schedule consistently, because only small proportion of event-based users reported adherence difficulties based on the electronic health record data from Kaiser Permanente[45].The event-based users will be randomly sampled from the non-daily PrEP users, each of whose PrEP assessment interval’s durations is longer or equal to the corresponding duration of sero-protection period under the 2-1-1 dosing schedule.
   * Regular non-daily users: we will assume two regular non-daily dosing schedules (5 pills per week & 4 pills per week) that are practiced among the remaining non-daily users with equal proportion (i.e., 25% of the entire non-daily user population). These two schedules mimic the observed habits of PrEP intake every other day or on T and S (Tuesday, Thursday, Saturday, Sunday)[30]. To maintain the assumption that all PrEP users finish their medication supply before refilling their next prescription, the type of non-daily dosing schedules assigned will depend on the length of the PrEP assessment interval relative to its medication supply interval. (A medication supply interval is defined as the interval in which the medication supply of a prescription fill is expected to last under a regular non-daily dosing schedule.) For example, the medication supply interval of a prescription fill for the 4-pills-per-week schedule last days, where denotes the medication days supply of the prescription fill. Therefore, the 4-pills-per-week schedule is randomly assigned to the regular non-daily users, each of whose PrEP assessment interval’s duration is at least days. The same logic may be applied to the selection of regular non-daily users who would allow the 5-pills-per-week schedule. For each prescription fill, whether or not a MSM used PrEP on a day will be sequentially simulated through the PrEP assessment interval, as a Bernoulli random variable where the probability parameter is the daily PrEP intake probability based on the type of non-daily schedule (i.e., p=0.57 for the 4-pills-per-week schedule; p=0.71 for the 5-pills-per-week schedule). The simulation of PrEP intake dates would stop after exhausting the PrEP supply for that interval. However, if the sum of simulated PrEP intake across the PrEP use interval were less than the medication days supply of that interval, we will randomly sample the remaining dates of PrEP intake from the dates that were not simulated as PrEP intake date in the first round.

For each sensitivity analysis iteration, the following procedures will be implemented to create an iteration-specific dataset that will be analyzed using the same GBTM procedures as in the base analysis : 1) the random selection of daily PrEP users, event-based PrEP users, regular non-daily PrEP users; 2) the random simulation of PrEP intake dates for daily users and regular non-daily users, and subsequent creation of bwPDC time series and conversion to PrEP sero-protection time series; 3) the creation of PrEP sero-protection time series for the event-based users. The number of iterations implemented will depend on our computational capacity. We will conduct an empirical trial where we will first implement five iterations and compare the resulting iteration-specific cluster structures (i.e., the number of discovered groups, the proportion of group membership, group-specific predicted trajectories) with that obtained in the base analysis.

#### 2.2.6.2 Misclassification bias of PrEP sero-protection

A limitation to this analysis is that we do not have access to prescription records outside of the collaborating pharmacy chain’s administrative database. Misclassification of PrEP sero-protection level could occur for study subjects who switched pharmacy during their follow-up period, because they would be deemed having suboptimal sero-protection by our data mechanism even though they may be adhering to PrEP medication prescribed from other pharmacy. To address such an information bias, we will conduct a sensitivity analysis where we perform the GBTM procedures on individuals who also filled other non-PrEP prescriptions at the collaborating pharmacy chain after their last prescription fill of PrEP. In addition, we will be clear about the characteristics of our study cohort, especially the distribution makeup of users’ health insurance plans, to ensure transparency around the transportability of the discovered PrEP persistence patterns.

### 2.2.7 Strengths and limitations

One primary strength of our study is that we utilize the longitudinal PrEP prescription fill data to analyze PrEP persistence trajectories over time, which address the knowledge gap in the time-varying PrEP use habits that could not be observed in measures that summarize PrEP adherence in relatively extended time intervals. Furthermore, our study aims to provide straight forward, clinically friendly interpretation of the discovered PrEP persistence trajectory patterns that are linked to time-varying risk of suboptimal sero-protection from PrEP. In addition, we took advantage of the healthcare administrative data from a U.S. national pharmacy chain and formulate PrEP persistence trajectories based on daily dosing schedule. Such a study design will allow our analysis to complement on the findings from a similar study where Sagaon-Teyssier et al used GBTM to group the trajectories of PrEP adherence probabilities among French and Canadian MSM who follow event-based dosing schedule in a PrEP clinical trial (ANRS IPERGAY).

Our study is not without limitations. One primary limitation is that the collaborating pharmacy chain does not collect data on actual PrEP use from their customers (i.e., our study subjects). Although we attempt to address the impact of uncertainty around their actual PrEP use habits on trajectory clustering by simulating various common dosing schedules in our sensitivity analyses, the simulated datasets may or may not represent the study subjects’ true PrEP use experiences. Furthermore, the number of dataset simulations for assessing the impact of non-daily dosing schedules may be limited by our computational capacity.

In the sensitivity analysis, we attempt to address the possible misclassification bias of PrEP sero-protection for users who switched to a different pharmacy chain during their follow-up period by restricting the trajectory clustering analysis on study subjects who continue to fill non-PrEP medications after their last prescription fill of PrEP. However, there could be residual misclassification among PrEP users who frequently switch back and forth between two or more pharmacies. Despites this possibility, we deem that the proportion of such users exist in our study population is ignorable.

1. *Diagnoses of HIV Infection in the United States and Dependent Areas, 2018 (Updated)*, in *HIV Surveillance Report*. 2018, U.S. Centers for Disease Control and Prevention: Atlanta, GA.

2. *Estimated HIV incidence and Prevalence in the United States 2014 - 2018*, in *HIV Surveillance Report - Supplemental Report*. 2019, U.S. Centers for Disease Control and Prevention: Atlanta, GA.

3. *Preexposure prophylaxis for the prevention of HIV infection in the United States—2021 update: a clinical practice guideline*. 2021, U.S. Public Health Service: Atlanta, GA.

4. Grant, R.M., et al., *Preexposure chemoprophylaxis for HIV prevention in men who have sex with men.* New England Journal of Medicine, 2010. **363**(27): p. 2587-2599.

5. Molina, J.-M., et al., *On-demand preexposure prophylaxis in men at high risk for HIV-1 infection.* N Engl J Med, 2015. **373**: p. 2237-2246.

6. Grant, R.M., et al., *Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study.* The Lancet. Infectious diseases, 2014. **14**(9): p. 820-829.

7. Chan, P.A., et al., *Retention in care outcomes for HIV pre-exposure prophylaxis implementation programmes among men who have sex with men in three US cities.* J Int AIDS Soc, 2016. **19**(1): p. 20903.

8. Montgomery, M.C., et al., *Adherence to Pre-Exposure Prophylaxis for HIV Prevention in a Clinical Setting.* PLoS One, 2016. **11**(6): p. e0157742.

9. Chan, P.A., et al., *Long-term retention in pre-exposure prophylaxis care among men who have sex with men and transgender women in the United States.* Journal of the International AIDS Society, 2019. **22**(8): p. e25385-e25385.

10. Doblecki-Lewis, S., et al., *Healthcare Access and PrEP Continuation in San Francisco and Miami After the US PrEP Demo Project.* Journal of acquired immune deficiency syndromes (1999), 2017. **74**(5): p. 531-538.

11. Huang, Y.-L.A., et al., *Persistence With Human Immunodeficiency Virus Pre-exposure Prophylaxis in the United States, 2012–2017.* Clinical Infectious Diseases, 2020. **72**(3): p. 379-385.

12. Coy, K.C., et al., *Persistence on HIV preexposure prophylaxis medication over a 2-year period among a national sample of 7148 PrEP users, United States, 2015 to 2017.* Journal of the International AIDS Society, 2019. **22**(2): p. e25252-e25252.

13. Rolle, C.-P., et al., *PrEP Implementation and Persistence in a County Health Department Setting in Atlanta, GA.* AIDS and Behavior, 2019. **23**(3): p. 296-303.

14. Uthman, O.A., et al., *Depression and adherence to antiretroviral therapy in low-, middle- and high-income countries: a systematic review and meta-analysis.* Curr HIV/AIDS Rep, 2014. **11**(3): p. 291-307.

15. Mehrotra, M.L., et al., *The Effect of Depressive Symptoms on Adherence to Daily Oral PrEP in Men who have Sex with Men and Transgender Women: A Marginal Structural Model Analysis of The iPrEx OLE Study.* AIDS and behavior, 2016. **20**(7): p. 1527-1534.

16. Cochran, S.D., V.M. Mays, and J.G. Sullivan, *Prevalence of mental disorders, psychological distress, and mental health services use among lesbian, gay, and bisexual adults in the United States.* J Consult Clin Psychol, 2003. **71**(1): p. 53-61.

17. Hatzenbuehler, M.L., et al., *Prospective associations between HIV-related stigma, transmission risk behaviors, and adverse mental health outcomes in men who have sex with men.* Ann Behav Med, 2011. **42**(2): p. 227-34.

18. Miller, K.M., *Assessment of the Impact of Attention Deficit Hyperactivity Disorder on Diabetes Management and Glycemic Control among Adolescents and Young Adults with Type 1 Diabetes*. 2015: University of South Florida.

19. Uysal, S., et al., *The effect of Attention Deficit and Hyperactivity Disorder on Antiretroviral Treatment and Appointment Adherence among adults living with HIV.* J Infect Dev Ctries, 2020. **14**(3): p. 304-311.

20. Kanny, D., et al., *Racial/Ethnic Disparities in HIV Preexposure Prophylaxis Among Men Who Have Sex with Men - 23 Urban Areas, 2017.* MMWR. Morbidity and mortality weekly report, 2019. **68**(37): p. 801-806.

21. Hojilla, J.C., et al., *Characterization of HIV Preexposure Prophylaxis Use Behaviors and HIV Incidence Among US Adults in an Integrated Health Care System.* JAMA Network Open, 2021. **4**(8): p. e2122692-e2122692.

22. Kim, B., et al., *Geographic Density and Uptake of Pre-exposure Prophylaxis (PrEP) Among Young Gay, Bisexual and Other Sexual Minority Men: A Global Positioning System (GPS) Study.* AIDS and Behavior, 2021. **25**(2): p. 155-164.

23. Mouhanna, F., et al., *Multilevel analysis of persistence on HIV Pre-exposure Prophylaxis (PrEP) among 232,196 PrEP users in the United States in 2012-2018*. 2019, The George Washington University.

24. Siegler, A.J., et al., *Developing a Web-Based Geolocated Directory of HIV Pre-Exposure Prophylaxis-Providing Clinics: The PrEP Locator Protocol and Operating Procedures.* JMIR Public Health Surveill, 2017. **3**(3): p. e58.

25. Grossberg, R., Y. Zhang, and R. Gross, *A time-to-prescription-refill measure of antiretroviral adherence predicted changes in viral load in HIV.* J Clin Epidemiol, 2004. **57**(10): p. 1107-10.

26. Choo, P.W., et al., *Validation of patient reports, automated pharmacy records, and pill counts with electronic monitoring of adherence to antihypertensive therapy.* Med Care, 1999. **37**(9): p. 846-57.

27. Smith, D., M. Van Handel, and J. Grey. *By race/ethnicity, blacks have highest number needing PrEP in the United States, 2015*. in *Conference on retroviruses and Opportunistic Infections*. 2018.

28. Baugher, A.R., et al., *Racial, Ethnic, and Gender Disparities in Awareness of Preexposure Prophylaxis Among HIV-Negative Heterosexually Active Adults at Increased Risk for HIV Infection—23 Urban Areas, United States, 2019.* Morbidity and Mortality Weekly Report, 2021. **70**(47): p. 1635.

29. Rawlings, K., D. Magnuson, and R.M. Giler, *Proportion Days Covered as a Measure of Adherence in US PrEP Subjects.* AIDS Research and Human Retroviruses, 2014. **30**(S1): p. A43-A44.

30. Sewell, W.C., et al., *Nondaily Use of HIV Preexposure Prophylaxis in a Large Online Survey of Primarily Men Who Have Sex With Men in the United States.* Journal of acquired immune deficiency syndromes (1999), 2020. **84**(2): p. 182-188.

31. Nagin, D.S., *Group-based trajectory modeling: an overview.* Handbook of quantitative criminology, 2010: p. 53-67.

32. Nagin, D., *Group-based modeling of development Harvard University Press.* Cambridge, Mass, 2005.

33. Nagin, D.S. and C.L. Odgers, *Group-based trajectory modeling in clinical research.* Annual review of clinical psychology, 2010. **6**: p. 109-138.

34. Walsh, C.A., et al., *Identifying Adherence Patterns Across Multiple Medications and Their Association With Health Outcomes in Older Community-Dwelling Adults With Multimorbidity.* Value in Health, 2020. **23**(8): p. 1063-1071.

35. Librero, J., G. Sanfélix-Gimeno, and S. Peiró, *Medication Adherence Patterns after Hospitalization for Coronary Heart Disease. A Population-Based Study Using Electronic Records and Group-Based Trajectory Models.* PLoS One, 2016. **11**(8): p. e0161381.

36. Van Ryzin, M.J., et al., *Identifying atypical cortisol patterns in young children: The benefits of group-based trajectory modeling.* Psychoneuroendocrinology, 2009. **34**(1): p. 50-61.

37. Niyonkuru, C., et al., *Group-based trajectory analysis applications for prognostic biomarker model development in severe TBI: a practical example.* J Neurotrauma, 2013. **30**(11): p. 938-45.

38. Sagaon-Teyssier, L., et al., *A group-based trajectory model for changes in pre-exposure prophylaxis and condom use among men who have sex with men participating in the ANRS IPERGAY Trial.* AIDS Patient Care and STDs, 2018. **32**(12): p. 495-510.

39. Jones, B.L., D.S. Nagin, and K. Roeder, *A SAS procedure based on mixture models for estimating developmental trajectories.* Sociological methods & research, 2001. **29**(3): p. 374-393.

40. Monti, S., et al., *Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data.* Machine learning, 2003. **52**(1): p. 91-118.

41. Șenbabaoğlu, Y., G. Michailidis, and J.Z. Li, *Critical limitations of consensus clustering in class discovery.* Scientific reports, 2014. **4**(1): p. 1-13.

42. Nagin, D.S. and C.L. Odgers, *Group-Based Trajectory Modeling (Nearly) Two Decades Later.* Journal of Quantitative Criminology, 2010. **26**(4): p. 445-453.

43. Roeder, K., K.G. Lynch, and D.S. Nagin, *Modeling uncertainty in latent class membership: A case study in criminology.* Journal of the American Statistical Association, 1999. **94**(447): p. 766-776.

44. Molina, J., et al., *Incidence of HIV-infection in the ANRS Prevenir study in Paris region with daily or on-demand PrEP with TDF/FTC.* Age (years), 2018. **36**: p. 30-44.

45. Hojilla, J.C., et al., *Early Adopters of Event-driven Human Immunodeficiency Virus Pre-exposure Prophylaxis in a Large Healthcare System in San Francisco.* Clinical Infectious Diseases, 2020. **71**(10): p. 2710-2712.

46. Molina, J.-M., et al., *Efficacy, safety, and effect on sexual behaviour of on-demand pre-exposure prophylaxis for HIV in men who have sex with men: an observational cohort study.* The lancet HIV, 2017. **4**(9): p. e402-e410.

47. Wall, K.M., R. Stephenson, and P.S. Sullivan, *Frequency of sexual activity with most recent male partner among young, Internet-using men who have sex with men in the United States.* Journal of homosexuality, 2013. **60**(10): p. 1520-1538.