Cover page for clinical trials.gov

Official Study Title: Socio-Structural Intervention to Improve Pre-Exposure Prophylaxis Services for Cisgender Women (PrEP CGW)

NCT number: pending

Date of Document: May 3, 2021

STATISTICAL DESIGN & POWER

Aim 1: Intervention Development. We will recruit n=10 key informants for in-depth interviews and up to n=60 CGW and medical providers to contribute to the focus groups (2-3 groups per population with groups of 8-10 participants) and theater testing.

Aim 2: Feasibility Trial. Given the pilot nature of this study, it is not appropriate to conduct a power analysis; We are not testing effectiveness during this stage of the trial.[1]

Measures & Analysis. Data collection will be ongoing in through the EMR; reach and initial effects data will be extracted for all CGW at three time points (baseline, intervention, and discontinuation). More detailed data in regard to the intervention strategies (provider training and peer navigation) will be collected on Adoption and Implementation (feasibility and acceptability) at the end of the intervention time period as described in **Table 1** below.

Table 1. Aim 2 Outcomes, Data Sources, and Time Points for Evaluation

Outcome	Data Source/ Measure	Time of Evaluation*
Reach (also shows feasibility of intervention)		
% of eligible CGW counseled on PrEP by providers	EMR	Year 3, Quarter 2 (End of Intervention)
% of CGW who complete HIV prevention discussion with peer navigator	Comparison of peers' patient log with EMR patient visit log	Year 3, Quarter 2
Measure of Initial Effects (note: no statistical testing for effectiveness in pilot phase; measures are for feasibility of data collection and standard deviation calculation)		
% change in perceived medical bias	7-item Discrimination in Medical Settings Scale	Year 2, Quarter 3 & Year 3 Quarters 1&2
% eligible CGW initiate PrEP	EMR	Year 3, Quarter 2
Of CGW who initiate PrEP, % adherence at 1 month	EMR, DBS**, & Pharmacy Data	Year 3, Quarter 2
Of CGW who initiate PrEP, >75% will persist ≥ 1 month	EMR & Pharmacy Data	Year 3, Quarter 2
Adoption		
% of providers who participate in synchronous or asynchronous training	Attendance log	Year 3, Quarter 2
% change in knowledge/awareness of medical bias and positive PrEP counseling techniques	Provider training module pre- & post-knowledge questionnaires	Year 2 Quarter 4- Year 3 Quarter 1
Implementation		
Perceived feasibility of implementing the intervention	Provider & Peer Likert scale survey & Provider & Peer in-depth interviews (n=10)	Year 3, Quarter 2
Acceptability: % CGW who receive provider counseling or peer session and report high satisfaction Satisfaction with training (synchronous vs. asynchronous) and outreach intervention among providers and peers	CGW; Likert scale survey to rate affective attitude, or feelings towards the intervention, burden, or perceived amount of effort to participate, and congruence with individual values[2] Provider & Peer Likert scale survey; Provider & Peer in-depth interviews (n=10)	Year 3, Quarters 1-2 (Post-clinical visit or peer session) Year 2 Quarter 4 - Year 3 Quarter 2 (Beginning, mid, & end of intervention) Year 3, Quarters 1-2
Organizational perspective on the intervention	Provider & Peer in-depth interviews (n=10)	Year 3, Quarters 1-2
Patient perspective on the intervention	Patient in-depth interviews (n=20)	Year 3, Quarters 1-2

*See ***Study Timeline*** in PHS

For percentage outcomes we will use the following "traffic light" benchmark system for feasibility and acceptability in Aim 2: 0-50%: failure of the current intervention design; 51-75%: major revisions required; 76%: no or minor revisions needed for a confirmatory trial.

Questionnaires. RAs will invite all CGW seen in FPPC to participate in a 5-minute tablet-based survey to assess acceptability of the provider and peer components of the intervention and intention to initiate PrEP; CGW seen virtually will be contacted by phone and/or email and invited to participate in an online survey. We anticipate that approximately 300 CGW will complete the survey. Upon survey completion, participants will be invited to provide their contact information to participate in virtual or in-person IDI. Additionally, to assess intermediate effects of the provider trainings, we plan to use the 7-item Discrimination in Medical Settings Scale[3], which has been validated in Black populations, to capture changes in patient-reported experiences before and after implementation of trainings to capture participant perceptions of bias.

Adoption by providers will be measured using attendance logs, with the goal of >75% of providers completing trainings. Additionally, we will evaluate the educational impact of the provider trainings with pre- & post-knowledge questionnaires for all training participants. **Implementation** outcomes focus on feasibility and acceptability, which will be evaluated both quantitatively and qualitatively. For quantitative measures we will use a validated, 8-item Likert scale to be completed by providers and peer navigators.[4] We will also use the Theoretical Framework for Acceptability, with seven domains on a Likert scale to assess ethicality, affective attitude, experienced burden, opportunity costs, perceived effectiveness, self-efficacy, and intervention coherence.[2]

Feasibility and Acceptability will also be assessed in qualitative, in-depth interviews with patients and providers. Transcripts will be uploaded into the qualitative data analysis software, *Atlas.ti*,[5,6] for the purposes of coding and analysis. We will first develop an *a priori* code book[6] that reflects key analytic concepts related to topical areas described above contained in the semi-structured interview guide. During the process of reading and coding of transcripts using this initial coding scheme, additional codes may be added to document emerging domains of interest. Qualitative analysis will proceed by first exploring broad patterns of perspectives and experiences. For FGDs, we will look for patterns of consensus and divergence within groups, and comparative perspectives between groups e.g., potential PrEP users and providers in terms of receptiveness for intervention plans for both the provider trainings and peer navigation components. For IDIs, we will assess similarities and differences between types of participants for patients and providers, respectively, (e.g., variation by demographics, level and type of engagement) and then across the groups to gain insight into how the intervention was received and variation in acceptability.

These qualitative and quantitative data will then be triangulated to refine the interventions for the future full-scale trial.

Quantitative data will be entered into REDCap for structured data storage and analyzed using Python/R language. We will conduct a descriptive summary of the outcomes described in the above Table 1. Where appropriate, we will test item scale reliability using the intraclass correlation coefficient (ICC). In surveys we will also collect data on demographics (e.g., race, age, education, income, partner status), behavioral risk for HIV (e.g., number of sexual partners, HIV status of partner, consistency of condom use), medical history (e.g., comorbidities, duration of risky behaviors), and lifestyle habits (e.g., smoking, alcohol consumption, substance abuse). We will also describe the demographic and clinical characteristics of CGW included in the sample, including means (standard deviations) or medians (interquartile ranges), and frequencies (percentages). Following descriptive statistics, we will compare the proportions of CGW who initiate PrEP and persist ≥ 1 month by key characteristics (e.g., age, partner status). We will use logistic regression to explore characteristics of those initiating PrEP and continuing past one month.

In order to better estimate treatment standard deviations, we will explore RDD analyses in preparation for our future effectiveness trial. RDD is a quasi-experimental, pretest-posttest design, it is commonly used for causal research studies, which involve distinct cutoff values that assign subjects into different conditions and compare different conditions' outcomes. In the full-scale effectiveness trial, we will isolate the treatment effect by designing a baseline period (months 6-9 in year 2) and intervention period (month 12 year 1 to month 3, year 3), and will determine the cutoff points before the data collection. The design allows an unbiased estimate of the treatment effect by minimizing confounding bias and assessing the magnitude of the discontinuity on outcomes at the cutoff point. We will also consider potential confounders of demographics factors including age, income, partner status in the analysis to achieve comparability and exchangeability.

Capperlleri et.al. conducted a power analysis using $\alpha = .025$ (one-tailed) and provided recommendations for the sample sizes for detecting different standardized effect sizes.[7] Their results show: 1) a sample of 1078 would have a power of .80 for detecting a small effect size, 2) a sample of 150 would have a power of .80 for detecting a medium effect size, 3) a sample of 68 would have a power of .80 for detecting a large effect size, and 4) a sample size of 109 would have power of .95 for detecting a large effect size. While we are not conducting effectiveness testing at this stage, our numbers should allow for calculation of standard deviations to prepare for our fully powered trial. Based on their study, we will recruit an adequately sample size of CGW participants for the preliminary RDD analysis (i.e., 460 CGW participants during the intervention period, an additional 345 CGW participants in the baseline, and 460 in the discontinuation period).

Dr. Duan will conduct analyses in PythonR[8]. All data reporting will follow the CONsolidated Standards of Reporting Trials (CONSORT) guidelines.[9]

1. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *Journal of Psychiatric Research*. 2011;45(5):626-629. doi:10.1016/j.jpsychires.2010.10.008

2. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Services Research*. 2017;17(1):88. doi:10.1186/s12913-017-2031-8

3. Peek ME, Nunez-Smith M, Drum M, Lewis TT. Adapting the everyday discrimination scale to medical settings: Reliability and validity testing in a sample of African American patients. *Ethnicity and Disease*. 2011;21(4):502-509. Accessed April 27, 2021. /pmc/articles/PMC3350778/

4. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implementation Science*. 2017;12(1):108. doi:10.1186/s13012-017-0635-3

5. ATLAS.ti: The Qualitative Data Analysis & Research Software. Accessed January 7, 2021. https://atlasti.com/

6. Maxwell JA, Miller B. Categorizing and Connecting Strategies in Qualitative Data Analysis. In: Hesse-Biber SN, Leavy P, eds. Handbook of Emergent Methods. New York: The Guilford Press; 2008: 461-77. - Google Search. Accessed January 7, 2021.

7. Cappelleri, J. C., Darlington, R. B., & Trochim, W. M. (1994). Power analysis of cutoff-based randomized clinical trials. Evaluation Review, 18(2), 141-152.

8. R: The R Project for Statistical Computing. Accessed January 7, 2021. https://www.r-project.org/

9. Consort - Welcome to the CONSORT Website. Accessed January 7, 2021. http://www.consort-statement.org/