

B. Collection of Information Employing Statistical Methods

The Midwest HIV Prevention and Pregnancy Planning Initiative (MHPPPI) aims to reduce HIV infections and increase pregnancy planning among women in high HIV prevalence communities in the Midwest through building providers' capacity to offer expanded HIV prevention and family planning options. This application seeks the approval to conduct:

- 1) Descriptive survey ("climate survey") to describe current knowledge, attitudes and behaviors of a convenience sample of Midwest medical providers concerning pregnancy planning for HIV-positive women or women in relationships with HIV-positive partners.
- 2) Qualitative study ("qualitative substudy") with patients (HIV+ persons of reproductive health age or persons in relationships with HIV-positive partners) to describe and document experiences with reproductive health, family planning and experiences in medical settings.

1. Respondent Universe and Sampling Methods

There are between 1500 to 2000 physicians with professional licenses having either a gynecologic, obstetric, infectious disease, preventive medicine, or family practice specialty in each of the eight states we are surveying (according to the active licenses maintained by each individual state board). Given this, our potential universe for the descriptive climate survey ranges from 12000-16000. While it is beyond the scope of MHPPPI to train every provider within these specialty areas, it is important that the potential universe include an estimate of every provider as that is who we want to survey to understand current practices. MHPPPI education and training efforts are estimated to reach approximately 1,900 providers. The climate survey will inform training content but will not evaluate the efficacy of the trainings. Ideally our survey would reach each potential participant, however, the licensing database does not contain contact information and there is no way to ensure the survey is distributed to each licensed physician. Therefore, we are enumerating a convenience sample universe based on distribution lists, professional memberships, and partner contacts. Our target sample size is 300 completed surveys. Given the potential universe size of 12000-6000, with a 90% confidence level and a 5% margin of error we need to enroll at least 264 participants, thus our goal is an overall sample size of 300 which will allow for potential incompletes. The average completion rate for online surveys is 24%, therefore we aim to distribute the survey link to at least 1250 unique potential participants (Fan, 2010; Sheehan, 2001). While we aim to distribute the survey to at least 1250 potential participants, it is unlikely every email will be delivered, opened and the recipient will take the eligibility screener. We estimate about 64% of the 1250 will complete the screener, this estimate is based on previous studies with similar methodology (Fan, 2010).

There was no sample size calculation for the qualitative substudy, rather, we consulted the literature on best practice for constructing a purposeful sampling frame. Creswell, 2007, suggests enrolling 20 to 30 individuals when using the grounded theory approach. Likewise, the recommendation from Charmaz, 2006, is from 20-30 with option for additional participants should theories go unconfirmed or new information emerge. We propose to sample 20 HIV+ female patients of reproductive age or patients with HIV-positive partners to participate in the substudy. This is the lower end of the sampling frame and was chosen to minimize burden to the population.

2. Procedures for the Collection of Information

The survey is cross-sectional and is designed to provide a description on the landscape/climate of physicians' knowledge, attitudes and behaviors concerning

reproductive health, pregnancy planning, HIV prevention and HIV medical care. We will repeat the survey to describe the landscape in year 3 of the study; however, the survey will not track individuals over time. The survey methodology was chosen to assess the landscape of reproductive health and pregnancy planning options for HIV+ women and women in high prevalence communities offered by medical providers.

Our hypothesis is that our training program may potentially shift the landscape and increase the overall competency (e.g., knowledge), attitude and potentially behavior of medical providers (HIV primary & reproductive health care), thus describing the differences and similarities between findings before and after the trainings are offered may provide insight into any changes in the overall landscape. As the observed changes may be due to various environmental factors (e.g., drug formulary, policy, HIV prevalence, etc.) and sampling bias, the results from this study will be considered in the context of the limitations of the study's design. Data gathered from this collection will not be generalizable to the entire provider population in the Midwest and any communications/publications of the data will include discussions of the under or over representation of particular groups (sampling bias) and non-response bias that may impact the results. As bias in estimates/results could lead to incorrect conclusions about providers and the reproductive health landscape in the Midwest, we will evaluate nonresponse bias and include the methods and results in all publications and reports (Lineback, 2010). To evaluate nonresponse in the sample we will compare early and late respondents on outcomes (attitudes, behavior) (Lineback, 2010). Late respondents will serve as proxies for non-respondents in analyses; if statistically significant differences arise between the two groups it may be an indication of response bias (Lineback, 2010). Advantages to online data collection include ability to sample a wide geographic area (though, due to the convenience sampling methods used in this project, the collection of information from the sampled providers in this study will not be representative of the geographic distribution of providers in the Midwest MHPPPI network), survey accessibility when participants are available (important for clinicians who may have unconventional schedules), and automated data collection (reduces researcher time) (Wright, 2005) Our chief aim in this collection is to add substantive descriptive information to the field to establish the need as well as the content of the trainings for medical providers.

A qualitative substudy will be conducted with patients. This substudy will help inform the development of the curriculum content for the trainings. AFC and its partner agencies have heard many anecdotal stories from providers and patients about their experiences with reproductive health options; the qualitative interviews will allow us to systematically document and record these stories. We will thematically code all data and disseminate results in peer-reviewed journals with open access.

We will enroll up to 20 patients in the qualitative sub study; we will stratify participants on the two eligibility criteria aiming to enroll equal numbers in each arm (10 HIV+ people of reproductive age; 10 HIV- people with HIV+ partners). Participants will be recruited via partner agencies and will be screened for eligibility by evaluation staff at AFC (see eligibility screener).

Data Management & Statistical Analysis

AFC does not anticipate receiving any identifiable health information as part of this project; however, investigators are certified to conduct human subjects research (CITI) and are trained in the Health Insurance Portability and Accountability Act (HIPAA)

compliance and will adhere to that by maintaining all data in compliance to standards. No identifying information will be collected or stored. No personal identifying information will be requested or stored for any participants in either study (qualitative interviews or climate survey).

Data analysis will be conducted by PI Johnson.

The outcome under investigation is factors associated with routinely discussing family planning with patients (coded dichotomously yes/no).

Univariate analysis will describe the sample of the climate survey including frequency tables for categorical variables and displays of median and mean values for continuous variables. Bivariate analysis will be used to detect any associations between categorical variables using chi-square and ANOVA for continuous variables. Log binomial regression will be used to estimate prevalence ratios (PRs) from this cross sectional study. Log binomial models use the log link function to connect the dichotomous outcome to the linear predictor. The limitation to this model is it may fail to converge. If this is the case a Poisson regression with robust variance estimator can be used. Data will be analyzed using Stata V 15.0.

The qualitative study data will be thematically coded by the evaluation team using a grounded theory approach. Through a series of open coding we will reach saturation by looking for instances within transcripts as well as probing within interviews. We seek to publish the results of the study to add to the literature on this topic.

Quality Control and Quality Assurance

Data from the online survey will be checked for consistency in skip patterns and completeness on a weekly basis. Transcripts from the qualitative interviews will be reviewed for accuracy prior to coding.

Dissemination of Results & Publication Policy

Project findings will be disseminated via presentations at local and national conferences, as well as within professional networks, and via peer-reviewed journal articles. Each participating agency will receive a summary of findings and suggestions for implementing curricula outputs. All dissemination will include information about how to access trainings and promotional materials. To engage non-academic and non-professional audiences, AFC will include outcomes in its training curriculum delivery; develop a downloadable webinar and podcast; and feature the outcomes on the agency's website. Limitations will be included in all publications, with attention to the limitations of a convenience sample, cross sectional data, online survey methodology, and qualitative methods. We will include information about generalizability and interpreting results including potential threats of bias, including information, nonresponse bias, and selection bias.

3. Methods to Maximize Response Rates and Deal with Nonresponse

For the climate survey, a list of potential participants will be enumerated, in addition, partner agencies will be asked to forward the survey link to their networks. Partner agencies will be asked to tally how many forwards they complete (in order to assess number of survey links distributed). The survey link will be monitored in real time to assess number of initiates and number of completes vs. incompletes. A limitation of online survey methodology is non-response bias, in that individuals who do not choose

to participate may be different in terms of outcomes or exposures of interest. The evaluation team will identify, reduce and evaluate nonresponse patterns and potential bias (Krenzke, 2005). The team will plan to achieve a high response rate via using partners to reach out to their networks, promoting the survey in national and regional newsletters and listservs as well as notifying potential respondents prior to engaging them with the survey link via an introduction email describing the MHPPI program and goals as well as a brief overview of the survey; a follow-up email will be sent to invite participants to screen for eligibility. We will monitor response patterns during data collection and will target regions and subgroups if we detect low response rates in any particular region or subgroup, this will be an effort to reduce non-response. To evaluate nonresponse in the sample post-collection, we will compare early and late respondents on outcomes (attitudes, behavior) (Lineback, 2010). Late respondents will serve as proxies for non-respondents in analyses; if statistically significant differences arise between the two groups it may be an indication of response bias (Lineback, 2010).

4. Tests of Procedures or Methods to be Undertaken

AFC pilot tested the survey tool with 4 experts in the field. This procedure was exempt for clearance and was used to refine survey questions. No further pilot testing will occur.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Designed data collection and will analyze data:

Amy Johnson

312-334-0978

Akjohnson@aidschicago.org

Data collection will occur online using the data collection platform Qualtrics. Surveys will be administered using an online platform called Qualtrics. Qualtrics is an Application Service Provider (ASP) with a platform for creating and distributing online surveys. The platform records response data and can perform reporting on the data. All services are hosted online and require no download to create or respond to surveys. Qualtrics protects its servers by high-end firewall systems with regular vulnerability scans performed. Qualtrics uses Transport Layer Security (TLS) encryption (also known as HTTPS) for all transmitted internet data. There is an ability to password-protect surveys and have unique ID links. All data at rest is encrypted.

References

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative research*. Sage Publications Ltd, London.

Creswell, J. W. (2007). *Qualitative enquiry and research design: Choosing among five approaches*.

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in human behavior*, 26(2), 132-139.

Krenzke, T., Van de Kerckhove, W., & Mohadjer, L. (2005). Identifying and Reducing Nonresponse Bias Throughout the Survey Process. In *ASA Proceedings of the Joint Statistical Meetings*.

Lineback, J. F., & Thompson, K. J. (2010). Conducting Nonresponse Bias Analysis for Business Surveys. In *Proceedings of the American Statistical Association, Section on Government Statistics*.

Sheehan, K. B. (2001). E-mail survey response rates: A review. *Journal of Computer-Mediated Communication*, 6(2), 0-0.

Krenzke, T., Van de Kerckhove, W., & Mohadjer, L. (2005). Identifying and Reducing Nonresponse Bias Throughout the Survey Process. In *ASA Proceedings of the Joint Statistical Meetings*.

Sheehan, K. B. (2001). E-mail survey response rates: A review. *Journal of Computer-Mediated Communication*, 6(2), 0-0.

Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, 10(3), 00-00.