

Video Health Study: Analysis Plan

U.S. Department of Health and Human Services
Office of Adolescent Health
Evaluation of “Plan A”
NCT Number NCT03238313

January 2019

The Policy & Research Group
8434 Oak Street
New Orleans, LA 70118
www.policyandresearch.com
504.865.1545



INTRODUCTION: STUDY CHANGES TIMELINE

OVERVIEW

On April 4th, 2019, The Policy & Research Group reviewed the registry page for the *Evaluation of “Plan A”* on clinicaltrials.gov and found several inaccuracies in language and content. Updates were made to the registry to reflect the study design documented in our Evaluation Abstract (see Appendix B),¹ submitted to the Office of Adolescent Health on September 18, 2017 and the Impact Analysis Plan (included below), submitted to OAH on January 31, 2019.

In an effort to be transparent about changes that have been made to the registry since its inception, we outline below substantive changes that have been made to the design of the study over the course of the implementation period. Changes are organized chronologically. Detail is provided on what the original content of the Clinical Trials registration included, what the revised content now indicates, and (when applicable) provides a rationale for the change.

JULY 2017

- Office of Adolescent Health (OAH) Teen Pregnancy Prevention 9TPP) Tier 2B grantees receive notice that the funding period for the grant has been shortened by two years, shifting the end date of the grant from June 30, 2020 to June 30, 2018

SEPTEMBER 2017

- Change in the target sample size from 2,250 to 1,770. After reviewing enrollment trends, in consult with the grant Evaluation Technical Liaison and Program Officer, the decision was made to revise enrollment estimates. Baseline data gathered from the current study were used to update power calculations in order to estimate the sample size needed to detect hypothesized programmatic effects; the revised enrollment target reflects the minimum sample required to detect small program effects (MDES = .12).

APRIL 2018

- Federal judge rules in favor of the OAH TPP Tier 2B grantees and funding is reinstated. Funding period will now end on June 30, 2020.

OCTOBER 2018

- Estimated Primary Completion Date changed from October 31, 2018 to May 31, 2019 to provide sufficient time to reach target sample size
- Estimated Study Completion Date changed from November 30, 2019 to June 30, 2020 to provide sufficient time to complete follow-up period

¹ Note that several details related to study implementation described in the September 2017 Evaluation Abstract have changed over the course of the study, but for the purposes of transparency and documentation we have included the original abstract in Appendix B.

1) Research Questions that Address Program Effectiveness on Behavioral Outcomes

a. Primary research questions

1. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported use of long-acting reversible contraceptives (LARC) three months after receiving the treatment?
2. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported times having sex without a condom three months after receiving the treatment?
3. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported receipt of sexually transmitted infection (STI) testing three months after receiving the treatment?

b. Secondary research questions

1. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported use of dual methods of protection (condom and prescription birth control use) during vaginal sex three months after receiving the treatment?
2. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' use of other effective contraceptive methods three months after receiving the treatment?

c. Exploratory research questions

Exploratory research questions will investigate mediating factors, subgroup effects, outcomes at nine months, and other exploratory outcomes.

2) Description of the Intervention and Counterfactual Condition

The Video Health Study (VHS) is a randomized controlled trial (RCT) in which eligible, consenting participants are randomly assigned to a treatment or control intervention at one of eight study sites in California. Sites have been organized into four regions for administrative and staffing purposes.

The intervention is offered to females, aged 18 or 19, who self-identify as either Latina or African American, are visiting a participating reproductive health center, have been deemed appropriate for the study by clinic staff with regards to physical and mental health capacity, are not knowingly pregnant and not trying to get pregnant, have not participated in other Office of Adolescent Health-funded Teen Pregnancy Prevention (OAH TPP) programs, consent to participate in the study, and are willing to complete three- and nine-month follow-up questionnaires.

The treatment condition, *Plan A*, is 23-minute video intervention that promotes effective contraceptive use, use of dual methods of protection (condom use and prescription birth control use), and HIV/STI testing. The video aims to develop sexual health intentions, knowledge, and self-efficacy for communicating with providers about different contraceptive options that have been proven effective, such as LARC. The ultimate goal of the video is to influence viewers' uptake of LARC, reduce unprotected sexual activity, and increase receipt of STI testing.

The control (counterfactual) condition, *The Toxic Life Cycle of a Cigarette*, is a 17-minute video that details the negative effects that cigarettes have on the environment and on people who manufacture and use cigarettes. The informational video uses both narration and interviews to educate viewers on the dangers that cigarettes pose. The video includes no sexual or reproductive health content. The videos are delivered to study participants on laptops in a private room or area of participating reproductive health clinics.

- a. **Intervention condition:** The intervention or treatment condition is a 23-minute video called *Plan A*, which was developed by Sentient Research and is based on the entertainment-education model. Entertainment-education is a process of creating and implementing an entertainment program to increase knowledge, change attitudes, and change behaviors among a target audience regarding a social or health issue.² *Plan A* was developed specifically for an 18 to 19-year-old African American or Latina female audience, and delivers messages about the use of effective sexual protection and contraceptive options.
 - i. **Intended program components:** *Plan A* is an individual-level, clinic-based video intervention. Through a series of three, inter-related, soap opera-style vignettes, five primary topics are addressed in the video: 1) risk perception for pregnancy, STIs, and HIV; 2) contraception options, with an emphasis on LARC; 3) condom use and partner negotiation skills; 4) the importance of regular STI and HIV testing; and 5) comfort discussing sexual history, HIV/STI testing, and contraception methods with a health provider. Between each vignette, there is a short, animated sequence to deliver information about intrauterine devices (IUDs) and the implant. Prior to playing the video, study staff provide the participant with a pen and paper and let her know that if she has any questions while watching the video, she may write them down and ask her health care provider about them during her visit. After the video concludes, participants are given a post-card sized handout of additional resources with comprehensive information on teen pregnancy, contraception, and STIs.

The video developers contend that Social Cognitive Theory (SCT), the Extended Elaboration Likelihood Model (E-ELM), and the Health Belief Model (HBM) provide theoretical grounding and justification for why the intervention could instigate behavior change. SCT can be practically applied to explain how people learn, can be motivated to change, and believe that they have the ability to change, by watching others. Viewers of the *Plan A* video learn by watching characters make positive contraceptive and sexual protection choices (who are thus rewarded) or negative choices (which lead to negative health outcomes). In this way, viewer's outcome expectations for making positive or negative behavior choices are influenced.

Another important theory applied to this intervention is the E-ELM, which is based on the idea that when viewers are engaged and entertained, they are less likely to counter-argue with the behavior-change messages presented. This in turn makes influencing their beliefs, attitudes, and behaviors more likely. The *Plan A* video was developed with an appealing storyline, high production values, and persuasive messages that aim to feel unobtrusive to the viewer. The E-ELM contends that when viewers are entertained, they are more receptive to the pro-health messages. According to E-ELM, it is also critical that viewers can identify with the characters. Identification leads to increased message absorption and acceptance of the values and messages portrayed in the program. The *Plan A* video responds to this by portraying African American and Latina actresses.

² i Singhal A, Cody MJ, Rogers EM, Sabido M. Entertainment-Education and Social Change: History, Research, and Practice. London, England: Lawrence Erlbaum Associates Publishers, London. 2004.

The HBM provides another hypothetical mechanism for how messages and strategies can be used within the entertainment-education model to bring about behavior change. According to the HBM, a health-related behavior change will occur if a person believes the following: 1) that he or she is at risk of a negative condition, 2) that the condition can be avoided, 3) that the person has a positive expectation that the avoidance strategy (i.e., new behavior) will prevent the negative condition, and 4) that the person believes that he or she can successfully complete the avoidance strategy. In addition, the perceived benefits of the behavior change must be understood as sufficiently desirable to overcome any perceived barriers to action. *Plan A* is intended to be delivered immediately prior to an individual's visit with a sexual and reproductive health provider, so the messages it delivers about sexual protection and contraceptive options are well-timed with a moment when the individual is already considering her risk for sexually transmitted infections and/or pregnancy.

- ii. **Intended program dosage:** The 23-minute *Plan A* video is intended to be delivered to each participant in one screening, immediately prior to their appointment with a sexual and reproductive health provider.
 - iii. **Intended program content:** *Plan A* is a theory-guided, video intervention designed to increase LARC use, increase condom use, and encourage HIV/STI testing in 18-19 year-old African American and Latina females. The intervention aims to motivate behavior change by providing viewers with relevant facts about contraceptive and sexual protection options to create awareness, demonstrating risk reduction strategies for unwanted pregnancy and HIV/STIs, and modeling effective communication with sexual partners and health care providers.
 - iv. **Intended program delivery:** The *Plan A* video is intended to be delivered within a sexual and reproductive health clinic setting immediately prior to a visit with a health care provider. The video is to be viewed in a private or semi-private location on a laptop or tablet. For this study, PRG has contracted with a reproductive health service delivery organization as an implementation partner. PRG and our implementation partner have strategically selected eight clinics in the Central California region, which predominantly serve the target population, as the sites for this study. Study coordinators have been hired, who are staff of the reproductive health service delivery organization, to facilitate delivery of the intervention to study participants who are randomized to the treatment group. All study coordinators have been trained by PRG on the research protocol. This training provides detailed information on how to conduct participant recruitment, eligibility screening, consent/assent, enrollment, and randomization procedures, as well as data collection, entry, and submission procedures.
- b. **Counterfactual condition:** The control condition is a 17-minute video called *The Toxic Life of a Cigarette*. It was also developed with the entertainment-education model in mind, with the purpose of delivering information about the negative effects of cigarettes on the environment and on the people who manufacture and use cigarettes as well as with the purpose of motivating behavior change around tobacco use. The video does not include any sexual or reproductive health content.

The researchers have opted to provide an alternative intervention for those participants who have been randomized into the control condition for methodological reasons. The offer of the video to the control group is an attempt to mitigate any awareness effects that could arise if individuals assigned to the control condition were to become aware that they did not receive any "treatment".

The control condition is similar to the treatment in terms of its brief dosage and video-based delivery format. This conformity with the treatment should reduce any confounds that might arise if the counterfactual experience was different enough in aspects aside from the intended informative and motivational treatment.

- i. **Intended program components:** *The Toxic Life Cycle of a Cigarette* is a 17-minute, individual-level, video-based health intervention.
- ii. **Intended program dosage:** The control video is intended to be delivered to each participant in one screening, immediately prior to their appointment with a sexual and reproductive health provider.
- iii. **Intended program content:** *The Toxic Life Cycle of a Cigarette* is an anti-smoking health information video that uses both narration and interviews to educate viewers on the dangers that cigarettes pose and encourage them to not use tobacco products.
- iv. **Intended program delivery:** *The Toxic Life Cycle of a Cigarette* is intended be delivered in health or science classes at the high school or junior high school level. However, for the purposes of maintaining equivalence in the type of experience treatment and control participants receive prior to their visit with a health care provider at participating clinics, we have chosen to use this video in a clinic-based setting.

3) Study Design

- a. **Sample formation:** To become enrolled in the VHS sample, adolescent females must appear at one of the eight clinics participating in the study and meet eligibility criteria.

Potential study participants are identified in one of two ways. With the first method, study coordinators review administrative/clinic data for existing upcoming appointments at their assigned study clinics to identify potential participants (female; 18 or 19; and, if known, African American and/or Latina). The study coordinators then call patients with existing appointments, 1-2 days prior to their appointments to conduct a pre-screening and determine whether or not the individual is interested in participating in the study. If the individual is interested, she is asked to arrive 1 to 1 ½ hours prior to her appointment time at the clinic to be fully screened and, if she meets the eligibility criteria, enrolled in the study.

The second method involves clinic staff working with our study coordinators to identify walk-in patients who may be eligible for participation. The procedures used at each clinic vary to some degree, but generally, clinic staff working at the front desk and study coordinators regularly review the clinic schedule throughout the day to identify any walk-in patients who are 18-19 years old and African American or Latina females. As time permits, the study coordinators meet with any women who meet this criterion and ask them if they are interested in participating in the study. Interested individuals are formally screened for eligibility, and eligible individuals are enrolled in the study.

To enroll a participant, the study coordinator must conduct a full eligibility screening and obtain consent. If the individual provides consent to participate, she is randomized into the treatment or control condition and considered enrolled in the study. This set of participants, who are randomized into the study and offered the opportunity to receive one of the two video health interventions, constitutes the full intent-to-treat (ITT) sample. The offer to receive the *Plan A*

video health intervention is the ITT treatment that we investigate in the primary and secondary research questions.

- i.* **Eligibility criteria for target population:** a number of criteria have been established for participation in the study.

To be eligible, participants must:

1. Consent to participate;
2. Be female;
3. Be 18 or 19 years old at the time of their enrollment, based on self-reported date of birth;³
4. Self-identify as African American and/or Latina;
5. Be deemed appropriate for the study with regards to physical and mental health by a clinician, clinic staff, or the study coordinator;

They must not:

6. Knowingly be pregnant;
7. Be trying to get pregnant;
8. Have enrolled in the SpeakOut study between November 2016 to October 2018⁴;
9. Be already enrolled in the study.

- ii.* **Purposeful Sampling:** Any individual who meets eligibility criteria is asked if she would like to hear more about the study. If she says yes, the study coordinator discusses the consent procedures dictated by the IRB to allow the individual to make an informed decision about study participation. If she consents and is therefore randomized into a condition, she is enrolled in the study and considered part of the ITT study sample.

b. Random assignment process

- i.* **Unit of randomization:** Random assignment occurs at the individual participant level.

- ii.* **Random assignment procedures:**

Standard Approach: Electronic random assignment

Electronic random assignment is conducted prior to the administration of the online baseline questionnaire. Random assignment blocks of varying sizes assign participants to the treatment or control condition at an equal (i.e., 1:1) assignment ratio. Prior to beginning study activities in a region, PRG produces an electronic randomization allocation list specific to each of the four regions with an existing algorithm available in Stata (random allocation command, ralloc). The allocation lists are produced by a senior analyst, password-protected, and stored on a secure PRG server. Study coordinators in each region are given a sequential list of unique five-digit study IDs that they assign in sequential order to eligible individuals. Random assignment occurs when the unique

³ An individual is determined by the study coordinator to have met the eligibility criteria if her calculated age (defined as the difference between her screening date and date of birth) makes her 18 or 19 years old. However, there are several instances when participants misreported their ages (stating that they were 18 or 19 years old at screening, when in fact their calculated ages indicated that they were truly 17 or 20 years old), and the study coordinators enrolled them into the study based on the inaccurately reported age. In such situations, if a participant's date of birth is within one month of the screening date, (indicating that she reached 18 years of age within one month of the screening date, or that she turned 20 years old within one month of the screening date), she will still be included in our final analytic sample. Any participant whose calculated age falls outside this window will be excluded from our sample.

⁴ The SpeakOut study is another OAH-funded TPP randomized controlled trial being implemented at reproductive health clinics in Central California. The purpose of this study is to determine whether delivery of SpeakOut, a behavioral intervention to increase social communication about long-acting reversible contraceptive (LARC) methods among adolescents, is associated with increased uptake of LARC methods among the social contacts of SpeakOut recipients.

study ID number is entered into a field on the initial page of the online platform used to administer the baseline questionnaire. It is at this point that the ID number is associated with an assignment condition (treatment or control) in the random allocation sequence. At each baseline administration (after eligibility has been confirmed), the study coordinator will type in the ID number that is next on the list (going in ascending numerical order) into the study ID field of the questionnaire; the questionnaire will then provide a message to the study coordinator that indicates the random assignment for that ID number. The study coordinator is trained to not allow the participant to see the screen that shows the random assignment message. While the participant is completing the baseline questionnaire, the study coordinator will record the participant's ID number and allocation into the *Enrollment Form*.

Contingency Approach: Envelope-based assignment

An alternative method that is employed only in instances when Wi-Fi or Internet service is down is envelope-based assignment. Prior to the start of the study at each site, and on an ongoing basis as resupply is needed, PRG prepares 'assignment envelopes' (50 are initially prepared for each region). PRG uses the same algorithm as used for the electronic-based assignment to create separate regional allocation lists for the envelope-based assignments (study IDs are recognizably unique from the electronic-based regional ID numbers and do not overlap with IDs used for electronic-based assignment). Each sealed security envelope has the region name and unique study ID recorded on the outside and contains a piece of paper indicating the assignment condition (*Plan A* or *The Toxic Life Cycle of a Cigarette*) from the allocation list. The message on the paper within the envelope will be masked in such a way that if participants see the message they will not know to which condition they are assigned (same procedure as used for electronic random assignment). PRG then places the envelopes (in ascending numerical order) in a box to be given to and managed by the study coordinator in each study region.

At each baseline administration that requires envelope-based assignment, the study coordinator will pick up the next envelope in the stack; the number written on the outside of the envelope will be the assigned ID number for that participant. The study coordinator enters this number into the 'study ID' field on the questionnaire (administered on paper) so that this number is associated with the participant's questionnaire data. Prior to the completion of the baseline questionnaire, the study coordinator will open the envelope and read the paper inside that indicates the condition to which the participant is assigned. This is when random assignment occurs – when the envelope is opened associating that individual (and ID number) with a treatment condition. As above, the blocking procedure (various sized blocks) will ensure an (almost) equal number of treatment and control assignments at each site. The study coordinator enters the study ID and assignment in the *Enrollment Form* while the participant completes the questionnaire.⁵

- iii. **Blocking procedures:** Blocking occurs at the regional level. Each of the four regions has been assigned a range of pre-randomized ID numbers (for electronic randomization) and a package of pre-randomized assignment envelopes (described in the subsection above). Participants are enrolled and randomized at the individual level within each region.

⁵ Each ID number and its corresponding intervention assignment is logged by the regional study coordinator on the participant's *Enrollment Form*. ID numbers and assignments from the *Enrollment Form* dataset are then matched to PRG's randomization allocation dataset so that we can monitor the integrity of the randomization process. This should ensure, at a minimum, that the condition a particular participant is assigned is the one that is indicated in the assignment records. This is to say that the ITT "point of offer" treatment should, at minimum, formally retain all the properties of random assignment even if a study coordinator wrongly administers the incorrect intervention.

- iv. **Probability of assignment to treatment group:** The probability of assignment to the treatment group is intended to be equal to the probability of assignment to the control condition; that is p (assignment to treatment) = .5.
- v. **Potential for crossover/contamination:** Given the brief nature of the intervention and counterfactual conditions – namely a 23-minute video for the treatment group and a 17-minute video for the control group – the potential for crossover or contamination is very limited. Crossover or contamination could occur if a study coordinator mistakenly administers the incorrect condition to an individual. However, study coordinators have received extensive training on study procedures. Additionally, when randomization occurs using electronic random assignment, the video to which the individual has been assigned is set up to automatically play, and thus does not rely upon the study coordinator selecting the correct video. The only situation in which crossover or contamination could potentially occur is if electronic random assignment is not available because of loss of network connection, and the study coordinator inadvertently shows the wrong video to a participant, but these circumstances are unlikely.
- c. **Consent process:** There is no difference in the consent process for the treatment or control groups. Evaluation consent is a condition of eligibility for the study, so no individual is randomized to a condition until after informed consent is obtained. As incentive to participate, a \$40 gift card is given to individuals who enroll in the study each time they complete a questionnaire (baseline, 3-month, and 9-month), and a \$5 gift card is given to individuals each time they respond to a request by study staff to update their contact information.⁶

After the individual has been screened and found eligible to participate in the study, the study coordinator goes through the *Participant Informed Consent Form* with her. This provides the individual with information about the study, outlines why she has been invited to participate, and addresses any questions that may arise. This process involves a paragraph-by-paragraph exploration of the consent form by the study coordinator and the female and constant “checking in” with the individual to be sure she fully understands the study requirements. At the end of this exercise, she is asked to provide consent if she wishes to participate in the study.

- d. **Data collection:** Data used for investigating both primary and secondary research questions are obtained from the *Participant Questionnaire* administered at baseline, three months post-baseline, and nine months post-baseline. The questionnaire is used to collect data on study participants’ self-reported contraceptive use and knowledge, sexual behavior and experiences, and intentions, thoughts, and feelings related to sexual behaviors. It is administered three times at the following time points:
 - a. Baseline – at the enrollment session just prior to the participant receiving their assigned intervention
 - b. Three months post-baseline – three months after the baseline questionnaire administration
 - c. Nine months post-baseline – nine months after the baseline questionnaire administration

While we collect data at three times over nine months, our analysis of primary and secondary research questions is concerned only with data gathered at baseline and three months post-

⁶ At the beginning of March 2017, the IRB approved a request to increase the gift card incentive offered to participants living in one region for completion of the baseline questionnaire from \$40 to \$60. This was done due to the higher cost of living in the region relative to the other study regions, and because the \$40 incentive amount seemed insufficient to attract participants to the study. In April 2017, the IRB also approved a request to increase the gift card incentive offered to participants who did not complete their follow-up questionnaires within two months of the follow-up window opening from \$40 to \$60. This was done to increase the likelihood of capturing data from harder-to-reach participants.

baseline.⁷ Exploratory research questions will investigate mediating factors, subgroup effects, outcomes at nine months, and other, exploratory outcomes. Study participants who are offered *Plan A* and *The Toxic Life Cycle of a Cigarette* videos receive the same questionnaire. The questionnaire contains 75 items and takes, on average, 18 minutes to complete. The instrument was constructed by PRG staff and is composed of items and scales that have been used in previous research on sexual behaviors and contraceptive use. The instrument was reviewed by health professionals and pilot-tested by young women with similar characteristics to our proposed study population. The questionnaire includes the same items at each time point and will measure the same constructs with identical measures at each administration.

There are no differences in data collection procedures for treatment and control groups. Data collection is conducted identically for both groups.

There are minor variations in procedures at the different data collection points. For all data collection points, we administer the questionnaire using a web-based (online) survey administration tool that has Audio Computer Assisted Self-Interview (ACASI) capabilities. The baseline online questionnaire is administered by clinic-based study coordinators, whereas, follow-up online questionnaires are administered by PRG research assistants.⁸ Research assistants make every attempt to collect outcome data as soon as possible after each data collection window (three-month follow-up and nine-month follow-up) opens; however, the data collection window remains open for four months to allow sufficient time for participants to complete their questionnaires. Any questionnaires completed after a data collection window closes will not be included in the final analytic sample.

At each data collection point, the study coordinator or research assistant identifies a quiet and private space for the participant to complete the questionnaire. During the enrollment session, the baseline questionnaire is completed in the clinic. Follow-up questionnaires are completed in locations that ensure participant confidentiality as well as convenience, such as the study coordinator or research assistant's office space, coffee shops, or other public locations like university buildings. On the computer, the study coordinator or research assistant selects the appropriate link to the type of questionnaire to be administered, enters the participant's unique participant ID number, and gives the study participant some brief instructions about taking the questionnaire. The script for the instructions was developed by PRG and emphasizes the importance of the participant's honesty in answering questions and the confidentiality of their responses. Study coordinators or research assistants are instructed to read it prior to each questionnaire administration. The study coordinator or research assistant also provides the participant with a calendar that they can reference when taking the questionnaire.

The study coordinator or research assistant then tells the participant to click "Next" on the computer screen when ready to begin and turns the study computer over to the participant to

⁷ With the assumption that we maintain low attrition and that the RCT is executed with integrity, we could approximate an un-biased estimate of the average treatment effect of *Plan A* by comparing differences in the means of our outcome variables reported by the treatment group with those reported by the control group. We could then provide a compelling response to our research question by testing the hypothesis that there is no difference between the two groups using straight-forward hypothesis testing statistics (t-test). However, we propose to use regression-adjusted means as the primary estimate of *Plan A* program effects to improve the precision of our estimates. Refer to subsection 4f below for a more detailed description of our proposed analytic approach.

⁸ Study coordinators are study staff who work for the VHS, but who are recruited and hired by the reproductive health service delivery organization. Study coordinators are responsible for recruiting and enrolling participants into the study. Research assistants are study staff who work for the VHS, but who are recruited and hired by PRG. Research assistants are responsible for maintaining contact with study participants after enrollment and for all follow-up procedures. This study management structure was put into place because of the need for individuals who are enrolling participants to have access to reproductive health clinic appointment records for pre-screening purposes. The policies of our implementation partner do not permit external staff to have access to this information.

complete the questionnaire. The study coordinator or research assistant is present in the room while the participant completes the questionnaire, but study staff are trained to minimize the impact of their presence and to clearly explain that the reason they are staying is to facilitate questionnaire administration.⁹ During the baseline questionnaire, the participant is instructed to let the study coordinator know when she has finished the survey, and the study coordinator clicks the “Submit” button. This is done so that the study coordinator can go through the video intervention script and provide a paper and pen to the participant prior to the video playing. Upon submitting the questionnaire, the participant is directed to the Wistia online video platform to view the assigned video. She is told to inform the research staff when the video is finished playing. During follow-up questionnaires, the participant is instructed to click the “Submit” button herself. In both instances, the regional research assistant receives an automatic email letting her know the questionnaire has been completed.

If, for some reason, the internet is not working when a questionnaire needs to be completed, or if the participant’s needs dictate that she takes the questionnaire on paper (e.g., because the participant needs the Spanish-language version of the questionnaire, or due to computer illiteracy), the study coordinator or research assistant follows the same procedures as above, but the study participant completes the questionnaire using a self-administered paper version. The study coordinator or research assistant enters the study participant’s unique ID on the paper questionnaire and gives her the questionnaire and a large envelope. At the end of the regular questionnaire instructions, the study coordinator or research assistant instructs the participant to put her completed questionnaire into the envelope and seal it when finished. The participant then hands it to the study coordinator or research assistant who writes the study ID number on the outside of the envelope and stores it in a locked file cabinet. Paper questionnaires are mailed via FedEx to PRG on a regular basis and entered by PRG staff into the *Participant Dataset*.

Research staff attempt to have all questionnaires administered in person; however, two weeks after the data collection window opens, if a participant is unwilling or unable to complete follow-up questionnaires in person, administration options offered to participants are for them to either complete it online using a survey link emailed to them and their own personal computer or tablet or complete it on paper using a questionnaire mailed to them. Three months after a follow-up window has opened, if research assistants have not been successful in getting the participant to complete the questionnaire online or with a mailed copy, the final option offered is to complete a shorter version of the questionnaire over the phone in an interview format with the participant. We will run sensitivity analyses that exclude participants who were surveyed by phone from our analytic sample and report substantive differences in the results section of the report. Study participants typically receive a \$40 incentive for each follow-up questionnaire completed; however, in April 2017 the IRB approved the study’s request to offer \$60 incentives, when deemed necessary, to participants who reach the third or fourth month of the data collection window and have not yet completed the follow-up questionnaire

e. Data collection related to additional analyses:

Recruitment Log

The *Recruitment Log* is an electronic database housed in the *Zoho Creator* application. Data from the paper *Eligibility Screening Form* are entered into the electronic *Recruitment Log* by the study coordinator. Each patient who is screened for the study is entered as one record in the log. The *Eligibility Pre-Screening Script* is a paper form used by study coordinators to recruit and ‘pre-screen’ over the phone potential study participants (females ages 18-19) with upcoming appointments at the clinics. The form includes a study introduction script and several pre-screen

⁹ There may be some occasions when the study coordinator needs to briefly leave the room, but in these instances, the study coordinators have been instructed to still remain available to participants.

questions to determine if a patient is potentially eligible and interested in enrolling in the study. The *Eligibility Screening Form* is a paper form used by study coordinators to collect eligibility information in person from potential study participants who come into the clinic for a reproductive health visit. The form includes a study introduction script, eligibility screening questions to be asked of the patient, and collects eligibility determination data. Data from these forms are entered into the *Recruitment Log* database by study coordinators on an ongoing basis.

Participant Database

The *Participant Database* is an electronic database housed in a *Zoho Creator* application. Data from the paper *Enrollment Form* are entered into the electronic *Participant Database* by study coordinators; data from the 3- and 9-Month *Follow-up Data Collection Forms* are entered into the database by research assistants. Each eligible patient enrolled in the study should have one record in this database for each of the three forms housed here (*Enrollment Form*, *3-Month Data Collection Form*, and *9-Month Data Collection Form*). The paper *Enrollment Form* collects administrative data, participant information on language and gift card preferences, data on adherence to the study procedures, questionnaire completion data, baseline incentive tracking numbers, and any notes on issues/changes to the study protocol. The paper 3- and 9-Month *Data Collection Form* forms are completed by the research assistant at the 3-month and 9-month follow-up data collection points. One form should be completed for each enrolled participant. The forms collect administrative participant information, questionnaire completion data, incentive tracking data, and notes on issues/concerns with the 3-month or 9-month questionnaire administration session.

Wistia video platform

The majority of study participants watch their assigned video on the Wistia online video platform. Wistia collects dosage data on the percent of the video played for each participant. Research staff at PRG download these data regularly to monitor program dosage. Data from the Wistia platform will be used to measure whether participants received the condition to which they were assigned and dosage received.

4) Analysis

a. Outcome measures:

Primary outcome measures

Our primary research questions ask to what extent the offer to watch *Plan A* relative to the offer to watch *The Toxic Life Cycle of a Cigarette* impacts participants' reported: 1) use of LARC; 2) times having sex without a condom; and 3) receipt of STI testing three months after receiving the treatment. We describe below the specific operationalization of each of these three primary outcome measures.

i. LARC Use

LARC use is constructed as a dichotomous variable – participants are either coded as currently using a LARC or not currently using a LARC. Data used to assess the impact of the treatment (*Plan A*) on LARC use are obtained from the following item on the *Participant Questionnaire*, which is administered to both the treatment and control groups at baseline and three months post-baseline.

- Which of the following methods of prescription birth control are you currently using?
 - None: I am not currently using any of these methods
 - Oral contraceptives (for example, the pill)

- *The patch (for example, Ortho Evra)*
- *The shot/injection (for example, Depo-Provera)*
- *The ring (for example, NuvaRing)*
- *The implant (for example, Implanon or Nexplanon)*
- *IUD (for example, ParaGard, Skyla, or Mirena)*

Persons who indicate that they are currently using either *The implant* or *IUD* are considered to be currently using LARC and are coded as 1. Persons who select *None*, *Oral contraceptives*, *The patch*, *The shot/injection*, or *The ring* are considered to not be currently using LARC and are coded as 0.

Plan A will be considered to have a positive impact on LARC use if, as compared to participants who are assigned to the control group, a larger proportion of participants who are offered *Plan A* report using LARC at the three-month follow-up and the difference between groups is statically significant.

ii. Times having sex without a condom

We operationalize times having sex without a condom as a risk outcome; that is, we measure the frequency with which participants engage in the risk behavior of having sex without a condom, rather than the frequency with which they engage in the safe sex practice of using condoms. Constructing the variable in this way allows us to examine the self-reported sexual behaviors of the full analytic sample of participants, regardless as to whether or not they are sexually active.

Specifically, times having sex without a condom is constructed as a continuous variable – the number of times in the past three months a participant *does not* use condoms while engaging in any type of sex.¹⁰ Data used to assess the impact of the treatment (*Plan A*) on condom use are obtained from the following three items on the *Participant Questionnaire*, which is administered to both the treatment and control groups at baseline and three months post-baseline.

- *In the past three months, how many times have you had vaginal sex without using a condom?*
- *In the past three months, how many times have you had oral sex without using a condom?*
- *In the past three months, how many times have you had anal sex without using a condom?*

Persons who indicate that they have not ever had a particular type of sex (vaginal, oral, or anal) or have not had that type of sex in the past three months are coded as having that type of sex without a condom zero times.¹¹ The final outcome measure is calculated by summing individuals' responses to these three items (the number of times they report they did not use condoms during vaginal, oral, and anal sex).

Plan A will be considered to have a positive impact on times having sex without a condom if, as compared to participants who are assigned to the control group, the number of times

¹⁰ *Plan A* is primarily aimed at reducing STI risk through the development of condom and STI knowledge. Since STIs can be transmitted through any type of sexual contact (i.e., vaginal, anal, or oral), our measure of times having sex without condom use is not limited to sexual intercourse but includes all self-reported sexual activity.

¹¹ The *Participant Questionnaire* contains sexual behavior questions that use a three-month recall period. As research has consistently found that memory of behaviors/events decreases over time and accuracy of recall is negatively associated with length of recall period (Clarke et al. 2008; Schwarz and Oyserman 2001), we use items with three-month recall periods to construct our measures of sexual behaviors since these should elicit more accurate responses than a longer recall period (e.g., six-month).

having sex without condoms reported by participants assigned to *Plan A* at the three-month follow-up is larger than the times reported by control participants and the difference between groups is statically significant.

iii. STI Testing

STI testing is constructed as a dichotomous variable – participants are either coded as having been tested for STIs in the past three months or not tested.¹² Data used to assess the impact of the treatment (*Plan A*) on STI testing are obtained from the following two items on the *Participant Questionnaire*, which is administered to both the treatment and control groups at baseline and three months post-baseline.

- The first question asks: *To the best of your knowledge, have you ever been tested for other sexually transmitted infections/diseases (STIs/STDs), such as chlamydia or gonorrhea?*
- If respondents select *Yes* to the first question, they are then asked this second question: *Have you been tested for STIs/STDs other than HIV in the past 3 months?*

Persons who select *Yes* to the second question are coded as 1, indicating that they have been tested for STIs/STDs other than HIV in the past 3 months. Persons who select *No* to the first and/or second questions are considered to not have been tested and coded as 0.

Plan A will be considered to have a positive impact on STI testing if, as compared to participants who are assigned to the control group, a larger proportion of participants who are offered *Plan A* report STI testing at the three-month follow-up and the difference between groups is statically significant.

Secondary outcome measures

Our secondary research questions ask to what extent the offer to watch *Plan A* relative to the offer to watch *The Toxic Life Cycle of a Cigarette* impacts participants' reported: 1) use of dual methods of protection (condom and prescription birth control use) during vaginal sex; and 2) use of other effective contraceptive methods (pill, patch, ring, or Depo-Provera) three months after receiving the treatment. We describe below the specific operationalization of each of these two secondary outcome measures.

i. Use of dual methods of protection

Use of dual methods of protection during vaginal sex is constructed as a proportionate continuous measure – the proportion of number of times in the past three months a participant has used both a condom and prescription birth control out of the total number of times reported having vaginal sex in the past three months.^{13,14} Data used to assess the impact of the treatment (*Plan A*) on dual methods of protection are obtained from the following two items on the *Participant Questionnaire*, which is administered to both the treatment and control groups at baseline and three months post-baseline.

- *In the past three months, how many times have you had vaginal sex?*
- *Please think about how many times you said you had vaginal sex in question 20. Of the times you had vaginal sex, how many times would you say you used*

¹² See footnote **Error! Bookmark not defined.**

¹³ Participants are provided with a box directly above this question, which lists the specific types of prescription birth control to which we are referring. This list includes the following: *birth control pills; the patch (for example, Ortho Evra); the shot (for example, Depo Provera); the ring (for example, NuvaRing); the implant (for example, Implanon or Nexplanon); and IUD (for example, ParaGard, Skyla, or Mirena).*

¹⁴ See footnote **Error! Bookmark not defined.**

dual methods of protections – that is, how many times did you use both a condom and one of the listed methods of prescription birth control at the same time?

Only those participants who report having vaginal sex in the past three months at both baseline and three-month follow-up will be included in analysis.. If an individual responds *No* to the question “*Have you ever had vaginal sex?*” or responds *No* to the question “*In the past 3 months, have you had vaginal sex, even once?*”, the individual will be excluded from analysis of this measure. Because of the fact that the mechanism defining the analytic sample for this research question is not random and there are theoretical reasons to believe that treatment assignment may affect inclusion (i.e., individuals who are offered the treatment may be more likely to abstain from sexual activity at follow-up and therefore excluded from the analytic sample) this constitutes an endogenous subgroup.¹⁵

The proportionate measure will be calculated as the number of times of reported dual methods of protection during vaginal sex in the past three months, divided by the number of times of reported vaginal sex in the past three months, with values ranging from 0 to 1.

Plan A will be considered to have a positive impact on use of dual methods of protection if, as compared to participants who are assigned to the control group, participants who are offered *Plan A* report using dual methods of protection more consistently (i.e., a larger proportion of the times they have vaginal sex) at the three-month follow-up and the difference between groups is statically significant.

ii. Use of other effective contraceptive methods

Use of other effective contraceptive methods is constructed as a dichotomous variable – participants are either coded as currently using other effective contraceptive methods or not currently using any contraceptive methods. Data used to assess the impact of the treatment (*Plan A*) on use of other effective contraceptive methods are obtained from the following item on the *Participant Questionnaire*, which is administered to both the treatment and control groups at baseline and three months post-baseline.

- *Which of the following methods of prescription birth control are you currently using?*
 - *None: I am not currently using any of these methods*
 - *Oral contraceptives (for example, the pill)*
 - *The patch (for example, Ortho Evra)*
 - *The shot/injection (for example, Depo-Provera)*
 - *The ring (for example, NuvaRing)*
 - *The implant (for example, Implanon or Nexplanon)*
 - *IUD (for example, ParaGard, Skyla, or Mirena)*

Persons who indicate that they are currently using either *Oral contraceptives*, *The patch*, *The shot/injection*, or *The ring* are considered to be currently using other effective contraceptive methods and are coded as 1. Persons who select *None* are considered to not be currently using other effective contraceptive methods and are coded as 0.

¹⁵ Endogenous subgroups are groups of participants that are defined by an outcome measure. The expected balance in observed and unobserved characteristics between treatment and control participants that results from random assignment cannot be guaranteed in such subgroups and thus effect estimates may be subject to bias. Tests of baseline balance will be conducted; if we have reason to believe there is imbalance between treatment and comparison groups for this outcome, we will reduce this to an exploratory research question.

It should be noted that this measure will only be constructed for individuals who select either *None*, *Oral contraceptives*, *The patch*, *The shot/injection*, or *The ring* to the question “Which of the following birth control methods are you currently using?”. If an individual selects *The implant* or *IUD* to this question, the individual will be excluded from analysis of this measure. As such, the subgroup of participants in which we conduct this analysis will be considered an endogenous subgroup (i.e. not inclusive of our full analytic sample) because it is constructed based on a participant’s response to an outcome measure.¹⁶

Plan A will be considered to have a positive impact on use of other effective contraceptive methods if, as compared to participants who are assigned to the control group, a larger proportion of participants who are offered *Plan A* report using other effective contraceptive methods at the three-month follow-up and the difference between groups is statically significant.

- b. **Analytic sample:** In Central California, eight clinics have been recruited to participate in the study. As explained in the *Study Design* section, clinic-based study coordinators review administrative/clinic data for patients with appointments at the study clinics to identify potentially eligible participants.

Potential participants are eligible if they: 1) are female; 2) are between the ages of 18 and 19 years of age; 3) are Hispanic and/or Latina; 4) are not knowingly pregnant or trying to get pregnant; 5) consent/assent to participate; 6) did not enroll in the SpeakOut study between November 2016 to October 2018¹⁷; and 7) are deemed appropriate for the study with regards to physical and mental health by a clinician, clinic staff, or the study coordinator.

After eligibility is determined, participants are randomized into the treatment (*Plan A*) or control (*The Toxic Life Cycle of a Cigarette*) condition according to procedures detailed in subsection 3b. The act of randomization into either the treatment or control arm constitutes the offer to participate in the intervention and is the point at which the individual becomes a participant in the study. The analytic sample is defined as all participants who were randomized into either the treatment or control conditions and who have reported sufficient outcome and covariate data.¹⁸ Missing data procedures are outlined in subsection 4cv below.

- c. **Data cleaning:** Prior to analysis, PRG staff will systematically screen or review the analytic variables (baseline and outcome) to identify invalid, inconsistent, outlying, missing, and unreliable data.¹⁹ New variables are created in which data that are deemed unusable (i.e., invalid

¹⁶ See footnote 15.

¹⁷ See footnote 4.

¹⁸ As outlined in subsection 4cv, our benchmark approach is to impute baseline/covariate data. As such, sufficient baseline/covariate data means all cases where data are not unit missing. We do not anticipate that we will have different analytic samples for our outcomes of interest; data are expected to be missing entirely for any given respondent at any observation point or not. If for whatever reason analytic samples are different for different outcomes, we will assess baseline equivalence separately for each analytic sample.

¹⁹ We propose to document the prevalence of inconsistent and missing data in a descriptives table presented as an appendix in our final impact report. Along with our presentation of sensitivity analyses, we will present tables that present the prevalence of unit and item missing (which result from nonresponse) as well as inconsistent, unreliable, and invalid data for both treatment and control samples. Regarding inconsistencies specifically, for each sexual behavior variable included in our model specifications (which could therefore influence the constitution of the analytic sample) we will include the following: sample size (the number of observations prior to recoding of inconsistencies) and the number of observations that are inconsistent over-time. If paper questionnaires lead to internally inconsistent data, we will also report on this.

or unreliable) are coded as missing and flagged according to missing data type; all other data are retained, unchanged.²⁰ The steps taken in this data cleaning process are outlined below.

- i. **Identify and flag unreliable cases:** The first step in the data screening process is to identify and flag entire cases (i.e., entire questionnaires) that are unreliable. By unreliable, we mean that we have sufficient reason to believe that the respondent's answers are not honest representations of their behaviors, knowledge, and beliefs. These cases are treated as missing and excluded from our benchmark analyses.

Cases are flagged as unreliable when responses follow a clear, deliberate pattern. This data cleaning procedure is informed by the data processing rules established for the National Survey on Drug Use and Drug Health (NSDUH) and for the Youth Risk Behavior Survey (YRBS), which treat records that follow defined patterns of responses as missing.²¹ PRG flags the following cases as unreliable: a) the same response option is chosen for all multiple choice questions; b) responses alternate between only two response options; or c) responses alternate systematically, starting with one response option, alternating through all options in order until exhausted then beginning again (in the same or in reverse order). If other response patterns are observed over the course of the evaluation, they will be added to PRG's list of unreliable response patterns.

Data for cases that are deemed unreliable are treated as *unit missing* and excluded from benchmark analyses. However, sensitivity analyses that include the unreliable data will be conducted and results will be reported in an appendix of the report.

- ii. **Identify and flag invalid responses:** The second step in the data screening process is to inspect the data for instances in which responses are invalid because they are outside of a pre-determined range of plausible or acceptable values. Each questionnaire type (e.g., baseline, three-month follow-up) has a codebook, which is prepared by a PRG staff, that contains variable names, valid variable values or ranges of values, and when applicable value labels.²² Referring to the codebook, a senior or lead research analyst performs diagnostics in STATA to ensure that responses to all analytic measures are valid (i.e., data are within ranges specified in the codebook). A data analyst inspects the data using two commands in STATA. First the analyst uses the command `sum variable_name`, which provides summary statistics (mean, minimum, maximum, standard deviation) for all numeric variables. The analyst checks that the minimum and maximum values are valid. If this command reveals there are values out of range, the analyst then inspects the data using the command, `tab variable_name, missing`, which provides a frequency table of all values (including missing values) so the analyst can identify and flag all values that are out of range as invalid and recode these values to missing (code as ".i").

Data that are recoded to missing are treated according to our missing data approach. Briefly, our benchmark approach is to adjust missing baseline data and include in analysis; we exclude observations with missing outcome data from analysis.

²⁰ A note on missing values: STATA provides a series of missing value codes that allow us to "flag" missing data according to why they are missing. Data that are missing due to unit nonresponse (a questionnaire was not completed) are coded using the "system missing" value (.). All other types of missing data are coded using "extended missing" values (e.g., ".a", ".b").

²¹ See *Comparing and Evaluating Youth Substance Use Estimates from the NSDUH and Other Surveys* retrieved February 27, 2018 from <https://www.samhsa.gov/data/sites/default/files/NSDUH-M9-Youth-2012/NSDUH-M9-Youth-2012.pdf>.

²² Regardless as to whether data are nominal, ordinal, or continuous, all response options are coded in STATA as numeric values; values are labeled according to corresponding category names when data are nominal or ordinal. As an example, the variable gender is a nominal variable; however, it is treated as a dummy variable where females are coded as "1" and males are coded as "0". The only acceptable values for this variable then are 0 and 1; any other values are out of range.

iii. **Identify and flag outliers:** The third step is to identify and flag severe outliers. Outliers (operationally defined below) are values that are extreme compared to other observations but are not plainly invalid. In the data cleaning process, we inspect outliers so that we can try to ascertain whether they are in fact true (or plausible) values or potentially a result of measurement error. The only variables for which we inspect outliers are those used in the construction of our outcome variables (times having vaginal sex, times having vaginal sex without a condom, times having oral sex, times having oral sex without a condom, times having anal sex, times having anal sex without a condom) because they have no upper limit (all other variables used in analysis are either categorical or have predicated upper and lower bounds). Our approach to identifying and flagging outliers is as follows.²³

- First, in STATA we use the *lv* (letter-value display) command to identify severe outliers. We define values as severe outliers according to their relation to the interquartile range (IQR). Severe outliers are defined as values outside of the *outer fences* of the population distribution.
 - $IQR = Q3(3^{rd} \text{ quartile or } 75^{th} \text{ percentile}) - Q1(1^{st} \text{ quartile or } 25^{th} \text{ percentile})$
 - Upper outer fence: $Q3 + 3 * IQR$
 - Lower outer fence: $Q1 - 3 * IQR$
- Second, we create an outlier indicator variable, where observations deemed severe outliers are coded as 1, all others are coded as 0.

Our benchmark analytic approach is to include data flagged as outliers in analysis, because we do not know for certain whether the values are true or invalid. However, we also run sensitivity analyses that exclude these data and report substantive differences in the results section of the report.

iv. **Identify and flag inconsistencies in reporting of sexual behaviors:** The fourth step in the data review process is to inspect the data and identify inconsistencies in sexual behavior outcome data. With repeated measures of sexual behaviors, two primary types of inconsistencies may occur – internal inconsistencies and over-time inconsistencies.²⁴ Internal inconsistencies refer to discrepancies in responses (to related questions) in the same survey administration. For instance, a respondent might say that she has not had sex in the past three months, but then indicates that she used condoms three of the times she had sex in the past three months. Over-time inconsistencies refer to instances in which lifetime reported behaviors decline or are completely recanted over time. For example, at baseline a respondent might say that she has had vaginal sex in her life, but on a subsequent administration of the survey she says she has never had vaginal sex.

In order to minimize internal inconsistencies in our primary and secondary outcomes, we built skip patterns into the online questionnaire – if participants indicate they have not had a particular type of sex in the past 3 months they are skipped out of more specific questions related to that type of sex. In addition, participants are precluded from indicating they had a particular type of sex without a condom more times than they said

²³ Rules for identifying outliers are informed by the following: Hamilton, Lawrence C. 2006. *Statistics with STATA: Updated for Version 9 and NIST/SEMATECH e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/>.

²⁴ Inconsistencies can occur for a number of reasons including social desirability bias and memory or recall issues on the part of the respondent and misunderstanding on the part of either the respondent or interviewer (Alexander et al 1993; Clarke, Fiebig, and Gerdtham 2008; Del Boca and Noll 2000; Harris et al 2008; Schroeder et al 2003; Schwarz and Oyserman 2001). These issues are especially common in self-reports of sexual behaviors where questions are of a sensitive nature and often respondents are asked to indicate the frequency and/or recency of behaviors over differing lengths of time (e.g., 30 days, 3 months, 6 months).

they had that type of sex. Because of this, no internal inconsistencies can exist in our primary and secondary outcome measures.²⁵

To address over-time inconsistencies, a research analyst examines all variables that are used to construct primary and secondary outcome measures, as well as any variables that may be used to logically impute values for primary and secondary outcome measures. If over-time inconsistencies are identified, both the baseline and follow-up values are flagged as inconsistent over time and recoded to missing (coded as “.k.”). Data that are recoded to missing are treated according to our missing data approach. Briefly, our benchmark approach is to adjust missing baseline data and include in analysis; we exclude observations with missing outcome data from analysis.

- v. **Missing data approach:** Assuming that our study design and procedures are sound, missing data pose perhaps the greatest threat to the internal validity of our RCT study and the ITT framework (Puma et al. 2009; Moher et al., 2010).²⁶ Randomization at the point of offer allows us to make causal statements about the effect of that offer because treatment and comparison samples are equal in expectation. For ITT framework to remain internally valid, however, the treatment and comparison groups must remain equal in expectation at the point of analysis. When the analytic sample is diminished by attrition or non-response, non-random differences (i.e., self-selecting) between the treatment and comparison groups may be introduced into the sample and estimates of program impacts may become biased. Although there is no consensus on how to resolve this, practical guidance on how to address and mitigate the problems associated with missing data have been published in education (Puma et al., 2009).

Our six-step decision processes for addressing this problem, as detailed below, are informed by this guidance. These steps articulate how we will deal with missing outcome or covariate data (that is variables outlined in the *Model specification and covariates* section and are necessary for the estimation of impacts). The benchmark approach that we have selected aims to mitigate the introduction of bias into our impact estimates, provide good estimates of uncertainty, and maximize the use of available data by adjusting missing baseline/covariate data. To test the robustness of this approach, and to verify these findings, we will report comparative findings using sensitivity analyses that also employ an alternative method which includes no adjustment (as outlined in step 6).

1. Using data cleaning procedures outlined in the *Data cleaning* section, identify inconsistent, outlying, unreliable, and invalid data in any analytic (i.e., outcome,

²⁵ In rare instances, paper-based questionnaires are completed by participants. This may occur if the study coordinator or research assistant is unable to open the online questionnaire because of network connection issues, or if the participant cannot be reached for an online administration and is instead mailed a paper questionnaire. Paper-based questionnaires include the same skip patterns as online questionnaires; however, unlike the online questionnaire where participants are skipped out of subsequent questions based on their response to previous questions and cannot report inconsistent times having different types of sex, it is possible for participants to provide responses to questions that they should not have answered or that do not make sense given their responses to previous questions. If and when such inconsistencies are observed in data originating from paper-based questionnaires, the participant's responses to the two or more items where inconsistencies are noted are recoded to system missing and will not be used in our primary and secondary outcome analysis.

²⁶ Puma, M.J., Olsen, R.B., Bell, S.H., Price, C. (2009). What to Do When Data Are Missing in Group Randomized Controlled Trials. (NCEE 2009-0049). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Moher, D. et al. (2010). CONSORT 2010 Explanation and Elaboration: Updated Guidelines for Reporting Parallel Group Randomised Trials. BMJ 2010;340:c869.

baseline, or covariate) variables, recode inconsistent and invalid data as missing, and flag unreliable and outlier data for analysis.²⁷

2. Report prevalence of unit and item missingness (which result from nonresponse) as well as inconsistent, unreliable, and invalid data for both treatment and control samples.²⁸
 3. Determine if logical imputations are possible for any analytic variables, other than outcome variables, that may have missing values (due to nonresponse) and logically impute where this is the case. We will not logically impute where the missing values are previously inconsistent, unreliable, or invalid.
 4. Determine if any individuals who are in the randomized sample have no data at all (i.e., unit missing) at baseline and at the three-month follow-up time point. If this is the case, our proposed benchmark approach is to use case deletion, as we feel it is the most straightforward and prudent approach for missing follow-up data recommended in Puma et al. (2009). These cases will be deleted from the analytic sample and attrition statistics will be reported.
 5. Determine if any individuals who are in the randomized sample (for each outcome) are missing baseline covariates or the baseline measure of the outcome variable. If this is the case, our proposed benchmark approach is to use dummy variable adjustment procedures, as we feel it is the most straightforward and prudent approach for missing baseline/covariate data recommended in Puma et al. (2009).
 6. Conduct sensitivity analyses by estimating results with missing baseline data excluded from the analysis (i.e., use case-wise deletion for all cases with missing baseline and outcome data). In an appendix, we will report our benchmark results next to the sensitivity analysis results to verify findings.
- d. **Assessment of baseline equivalence:** Baseline equivalence will be reported for all baseline measures of each outcome variable as well as relevant demographic and sexual behavioral measures. We first list and describe the measures we will use to examine the equivalence of our treatment and control group at baseline. After we identify the measures, we provide details on the diagnostic methods that we will use to assess any baseline difference that may exist between the treatment and control groups in the measures outlined below.

Demographic and Sexual Behavior Measures

Baseline equivalence will be assessed for four demographic variables and one baseline measure of sexual behavior (identified below). Each of the variables are constructed from participant self-responses to questions in either the *Eligibility Screening Form* or VHS baseline *Participant Questionnaire*. For the race variables, categorical responses to a single question are used to create multiple dichotomous variables. The sexual behavioral variable is constructed from participants' responses to three questions in the baseline *Participant Questionnaire*. We provide details on

²⁷ We will code missing responses with a unique missing code that identifies or flags these missing values according to the reason they are missing (i.e., nonresponse, invalid, inconsistent). Unreliable data are not recoded to missing, rather cases deemed unreliable are coded as 1 in an indicator variable, treated as unit missing, and excluded from analysis. See the *Data cleaning* section or Table 3 in Appendix A for details on how missing data are coded.

²⁸ For item missing values, we will only report prevalence of missing and inconsistent data for variables that are included in our model specifications and could therefore influence the constitution of the analytic sample.

variable coding below; details on variable construction can be found in Table 2 in Appendix A.

Demographic:

- Age at screening (continuous; range 18-19)²⁹
- Identify as Black/African American at screening (0 = identify as another race/do not identify race; 1 = identify as Black/African American)³⁰
- Identify as Hispanic, Latino, or of Spanish origin at screening (0=do not identify as Hispanic, Latino, or of Spanish origin; 1= identify as Hispanic, Latino, or of Spanish origin)³¹
- Completed high school (0 = has not completed high school; 1= has completed high school)

Sexual Behavior:

- Ever had sex (0=never had vaginal, anal, or oral sex; 1=has had vaginal, anal, and/or oral sex in lifetime);

Baseline Outcome Measures

In addition to the demographic and sexual behavior measures, we will assess baseline equivalency of the baseline observations of the outcome measures. We provide details on variable coding below; details on variable construction can be found in Table 2 in Appendix A.

- *Times having sex without condom in the past 3 months* at baseline (continuous; values range 0 to k , where 0= has had sex without condoms 0 times in past 3 months and k = number of times having sex without condoms in past 3 months)
- *Current LARC use* at baseline (0= not currently using a LARC; 1= currently using a LARC)
- *STI Testing in the past 3 months* at baseline (0=not tested for STIs/STDs in past 3 months; 1= tested for STIs/STDs in past 3 months)

Balance Assessment Methods

We propose to assess baseline equivalence of the treatment and control groups according to a multi-step procedure. Baseline equivalence statistics will be produced for each analytic sample.³² Only participants who provide baseline data to an outcome measure will be included in the analytic sample for that same outcome measure; thus the analytic sample used for each research question may vary slightly because of the exclusion of non-responders. As required by the “Identifying Programs that Impact Teen Pregnancy, Sexually Transmitted Infections, and Associated Sexual Risk Behaviors” review protocol, we will report the adjusted means and p-values of the differences in the baseline variable of interest for the treatment and control groups.³³ We will also report the standardized mean difference of each baseline variable for the treatment and control groups. This last statistic is not required by the review protocol but it is more consistent with the literature on balance statistics.³⁴

²⁹ Age at screening is determined using the participant’s self-reported date of birth.

³⁰ At screening, participants are asked “What is your race and ethnicity?” and provided with a list of the following response options: *White; Black or African American; Hispanic, Latino, or Spanish origin; American Indian or Alaskan Native; Asian; Native Hawaiian/Other Pacific Islander; Unknown; or Some other race/ethnicity*. Participants can select more than one category and they can also specify some other race/ethnicity. To be eligible for the study, participants must select either *Black/African American* or *Hispanic, Latino, or Spanish origin*.

³¹ See footnote 30.

³² Due to item missing outcome data, we expect there may be slight differences in analytic samples for each research question.

³³ Goesling, B., & Trenholm, C. (2016). Identifying Programs that Impact Teen Pregnancy, Sexually Transmitted Infections, and Associated Sexual Risk Behaviors. Mathematica Policy Research.

³⁴ The literature on balance statistics argues that significance testing is inappropriate for this diagnostic task (Austin, 2007; Imai et al., 2008; Austin, 2009; Stuart, 2010). Hypothesis tests can be misleading diagnostic measures of baseline equivalence because they conflate balance with statistical power

To establish baseline equivalence, we propose to generate model-based point estimates of the difference between the treatment and control groups for the identified baseline equivalence variables. We will report the adjusted means and p-values of the differences in the baseline variable of interest for the treatment and control groups. We will then compute the pooled standard deviation of these variables. Finally, we will produce a standardized difference of means by dividing the first term by the second.³⁵

step 1. First, we generate a model-based estimate of the difference between treatment and comparison groups on the pre-intervention measures identified above. Separate models will be run for each of the variables. The empirical model will be estimated with OLS (using Stata). If the measure is dichotomous we propose to use a linear probability model to estimate the predicted probability of group membership. The model is a reduced-form variation of the model that we use to estimate program impact (as detailed in the *Model specification and covariates* section, below).³⁶

$$Y_{baseline} = \beta_0 + \beta_1 T + \sum (\beta_p X_p) + \varepsilon$$

where:

$Y_{baseline}$ – is the baseline measure of the variable that we use to establish baseline equivalency (identified in Appendix A – Table 2). This variable is included as a covariate in the analytic model (see Table 2 for details on variable coding). Separate models will be estimated for each baseline equivalency measure specified above.

T – A dummy treatment indicator variable whose value equals 1 if the participant is randomized into the treatment group and zero otherwise.

X – Region (blocking variable) – An $n-1$ vector of region indicator dummy variables that are coded one if the intervention was delivered at region n and coded zero otherwise.

β_0 – The intercept term, which represents the adjusted mean value of the baseline equivalency measure for participants in the control sample, with all other variables in the model held constant at zero.

β_1 – This represents the adjusted (but not standardized) mean difference in the baseline equivalency variable between treatment and control participants.

ε – The residual or random variation that remains for each observation after the structural components of the model are estimated. It is the difference between the observed and the predicted values at the individual level.

step 2. Report the adjusted means and p-values of the differences in the baseline variable of interest for the treatment and control groups.

step 3. If the pre-intervention measure is continuous, we propose to use the following formula to calculate the pooled within-group standard deviation of the outcome measure:

$$S_p = \sqrt{\frac{(n_t - 1)S_t^2 + (n_c - 1)S_c^2}{(n_t + n_c - 2)}}$$

³⁵ Note that we will produce diagnostic estimates of baseline equivalence on the exact same samples of observations that we will use in our primary, secondary, and exploratory analyses. In other words, we will apply the missing data approach outlined in subsection 4cv prior to producing estimates of baseline equivalency on the pre-intervention measures. To validate this benchmark approach, we will also produce a sensitivity analysis that produces estimates of baseline equivalency without imputations (i.e., with all missing data). Benchmark and sensitivity estimates of baseline equivalency will be presented beside each other in an appendix.

³⁶ It is a reduced-form because individual-level, demographic covariates are omitted. It is a variation because the dependent variable is the baseline equivalence variable, not the outcome measure.

where: n_t and n_c are the sample sizes, and S_t and S_c are the participant-level standard deviations for the pre-intervention measures for the analytic treatment and comparison groups, respectively. We will produce separate calculations of the pooled standardized deviation for each variable used to establish baseline equivalence (as noted above).

step 4. Produce the standardized difference of means. If the pre-intervention measure is continuous, we will use Hedges' g as the formula to compute the standardized difference of means for the treatment and comparison groups:

$$g = \frac{\beta_1}{S_p}$$

Where: β_1 is the adjusted mean difference in the variable selected to establish baseline equivalence for the treatment and comparison groups (calculated in Step 1), and S_p is the pooled standard deviation (produced in Step 2).

For dichotomous baseline variables we will use the Cox index, which yields effect size values similar to the values of Hedges' g that one would obtain if group means, standard deviations, and sample sizes were available, assuming the dichotomous outcome measure is based on any underlying normal distribution.” Following this guidance, we propose to use the Cox index to estimate baseline equivalence for dichotomous baseline covariates. The formula is as follows:

$$d_{Cox} = \left[\ln\left(\frac{p_t}{1-p_t}\right) - \ln\left(\frac{p_c}{1-p_c}\right) \right] / 1.65$$

Where: p_t and p_c represent the probability of occurrence of the event (or characteristic) within the treatment and comparison groups, respectively.

- e. **Condition crossover and contamination:** Crossover will be defined as study participants assigned to the treatment condition who received any amount of the control video.³⁷ This will be determined from *Wistia* video platform data. We will calculate crossover using the following formula:

$$Crossover_{participant\ T} = \left(\frac{Not\ Received_{participant\ T}}{Base_{participant\ T}} \right)$$

where:

$Crossover_{participant\ T}$ - the proportion of participants randomly assigned to the treatment group who received the control video

$Base_{participant\ T}$ - the number of participants randomly assigned to the treatment group

$Not\ Received_{participant\ T}$ - the number of participants randomly assigned to the treatment group who received the control video

³⁷ In the OAH Impact Analysis Plan guidance for Cohort 2 Tier 2B grantees, crossover is described as occurring “when individuals randomly assigned to the intervention or counterfactual conditions are later found to be receiving the services intended to be offered to the other condition.” Given this, we calculate crossover in our sample as participants assigned to the treatment condition who received any amount of the control video, as this is the intended counterfactual condition.

Contamination will be defined as study participants assigned to the control condition who received any amount of the *Plan A* video.³⁸ This will be determined from *Wistia* video platform data. We will calculate contamination using the following formula:

$$Contamination_{participant\ C} = \left(\frac{Received_{participant\ C}}{Base_{participant\ C}} \right)$$

where:

$Contamination_{participant\ T}$ - the proportion of participants randomly assigned to the control group who received the *Plan A* video

$Base_{participant\ T}$ - the number of participants randomly assigned to the control group

$Received_{participant\ T}$ - the number of participants randomly assigned to the control group who received the *Plan A* video

Levels of crossover and contamination will be reported in the findings section of our final impact report.

- f. **Analytic approach for primary research questions:** As detailed in our primary research questions, this study investigates whether offering *Plan A* to participants impacts their reported LARC use, number of times having sex without condoms, and STI testing. We do this within the intent to treat (ITT) framework, which does not measure the effect of the participant's exposure to the treatment itself but rather the effect of the offer of the treatment relative to the offer of receiving the control condition. This framework maintains the integrity of the experimental structure by including all participants who were randomized (except those who attrite) in the analytic sample, thereby maintaining an exogenous assignment of participants to experimental condition. Bias can be insinuated, however, through self-selection if any participant who is randomized fails to provide outcome data.
- i. **Model specification and covariates:** The primary research questions under investigation in this study are whether offering *Plan A* to participants impacts their: (1) reported use of LARC, (2) times having sex without condoms, and (3) reported STI testing partners (see Table 1 in Appendix A for variable constructions). We propose to estimate these impacts using a regression that will model intervention effects as a function of assignment to *Plan A* (i.e., Treatment), relevant baseline covariates, a baseline measure of the outcome variable, and regional (blocking) indicators (see Table 2 in Appendix A for variable constructions).³⁹ Although a straight difference-of-means approach should provide unbiased estimates of the

³⁸ In the OAH Impact Analysis Plan guidance for Cohort 2 Tier 2B grantees, contamination is described as occurring "*when individuals assigned to the counterfactual condition end up receiving all or portions of the conditions intended only as part of the intervention.*" Given this, we calculate contamination in our sample as participants assigned to the control condition who received any amount of the *Plan A* video, as this is the intended treatment condition.

³⁹ With the assumption that we maintain low attrition and differential attrition and that the study otherwise executes the RCT with integrity, we should be able to estimate an un-biased estimate of the average treatment effect of the intent to treat participants with *Plan A* by comparing differences in the means of the outcome variable reported by the treatment group with those reported by the control group. We could then provide a compelling response to our research question by testing the hypothesis that there is no difference between the two groups using straight-forward hypothesis testing statistics (t-test). With that said, we propose a regression-based model that includes covariates, because randomization should ensure covariates are uncorrelated with the treatment variable (i.e., they should not affect the estimate of the treatment effect), and in the instance they are significant predictors of the outcome, their inclusion in a regression model will decrease the standard error of the estimates, making them more precise. See: Angrist, J. D., & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press; Rosenblum, M. and van der Laan, M. J. (2009), *Using Regression Models to Analyze Randomized Trials: Asymptotically Valid Hypothesis Tests Despite Incorrectly Specified Models*. *Biometrics*, 65: 937-945. doi:10.1111/j.1541-0420.2008.01177.x.

effect of the treatment, we propose a model-based approach because it will increase the precision of those estimates. The empirical model will be estimated with an OLS regression (using Stata).⁴⁰ We present the empirical model here:

$$Y_{Post} = \beta_0 + \beta_1 T + \beta_2 Y_{Pre} + \sum (\beta_p X_p) + \varepsilon$$

Where:

Y_{Post} – The outcome variable of interest, either: 1) times having sex without condoms in the past 3 months (continuous; values range 0 to k , where 0= has had sex without condoms 0 times in past 3 months, and k = number of times having sex without condoms in past 3 months); 2) current LARC use (0= not currently using a LARC; 1= currently using a LARC); or 3) STI testing in the past 3 months (0=not tested for STIs/STDs in past 3 months; 1= tested for STIs/STDs in past 3 months) reported by participant i at the 3-month follow-up. (see Table 1 for full details on the variable construction).

Y_{Pre} – The baseline measure of the outcome variable of interest reported by participant i at baseline (see Table 2 for full details on the variable construction); variable will be re-centered at the grand mean for analysis.

T – A dummy treatment indicator variable whose value equals 1 if the participant is randomized into the treatment group and zero otherwise.

X – A p vector of baseline (i.e., measured prior to receiving intervention or exogenous to treatment) participant-level covariates as well as blocking variables to account for the variation in outcomes associated with these groups. These covariates, listed in detail in Table 2 in Appendix A, will include:

- a) Age – self-reported age (based on date of birth) at screening (continuous; range 18-19); variable will be re-centered at the grand mean for analysis.
- b) Race – self-reported as Black/African American at screening (0 = identify as another race/do not identify race; 1 = identify as Black/African American); each of the variables will be re-centered at the grand mean for analysis.
- c) Ethnicity – self-reported as Hispanic, Latino, or of Spanish origin at screening (0=do not identify as Hispanic/Latino/Spanish origin; 1=identify as Hispanic, Latino, of Spanish origin); variable will be re-centered at the grand mean for analysis.
- d) High school education – self-reported education at baseline. A dummy variable (0= has not completed high school; 1= has completed high school); variable will be re-centered at the grand mean for analysis.
- e) Region – An $n-1$ vector of region indicator dummy variables that are coded one if the intervention was delivered at region n and coded zero otherwise. Region 1 is the reference category and is excluded from analysis. The dummy variables will be mean-centered for analysis to facilitate interpretation.⁴¹

β_0 – The intercept term, which represents, depending on the outcome measure of interest in the analysis, the outcome for the average control participant with all other variables in the model held constant at their mean.

β_1 – This is the parameter estimate of substantive interest. β_1 represents, depending on the outcome measure of interest in the analysis, either: 1) the adjusted mean difference in

⁴⁰ If the outcome measure is dichotomous, we will also construct a linear probability model to confirm our benchmark results. Beyond this, as part of our sensitivity analyses, we will construct a logistic regression model to explore any potential differences in our effect estimates given the model utilized.

⁴¹ Sites currently participating in the study have been organized into four regions for administrative and staffing purposes.

treatment and control participants' self-reported times having sex without condoms in the past three months at the three-month follow-up; 2) the adjusted odds ratio comparing treatment participants' current LARC use to control participants' use at the three-month follow-up; or 3) the adjusted odds ratio comparing treatment participants' STI testing in the past three months to control participants at the three-month follow-up.

ε – The error term or unexplained individual-level variance that remains for each observation after the structural components of the model are estimated. It is the difference between the observed and the predicted values at the individual level.

We will report model-estimated effects and the results of significance tests in the findings section of the final impact report. Statistical significance will be based on test statistics produced by Stata for the coefficient β_1 using a two-tailed test, with $p < .05$.

- ii. **Sample attrition:** Overall and differential attrition will be calculated using the full sample of participants enrolled in the study. This will be determined using data within the *Participant Dataset*. We will calculate overall attrition using the following formula:

$$Attrition_{participant} = 1 - \left(\frac{Assessed_{participant}}{Base_{participant}} \right)$$

where:

$Attrition_{participant}$ - the proportion of participants enrolled in the study who did not complete a 3-month follow-up questionnaire

$Base_{participant}$ - the number of participants enrolled into the study

$Assessed_{participant}$ - the number of participants who completed a 3-month follow-up questionnaire

Differential attrition will be calculated using the following formulas:

$$Attrition_{participant\ C} = 1 - \left(\frac{Assessed_{participant\ C}}{Base_{participant\ C}} \right)$$

$$Attrition_{participant\ T} = 1 - \left(\frac{Assessed_{participant\ T}}{Base_{participant\ T}} \right)$$

$$Differential\ Attrition_{t-c} = abs(Attrition_{participant\ T} - Attrition_{participant\ C})$$

Where:

$Differential\ Attrition_{t-c}$ - the absolute difference between the proportion of treatment group participants who did not complete a 3-month follow-up questionnaire and the proportion of control group participants who did not complete a 3-month follow-up questionnaire

$Attrition_{participant\ C}$ - the proportion of participants enrolled in the study and randomly assigned to the control group who did not complete a 3-month follow-up questionnaire

$Base_{participant\ C}$ - the number of participants enrolled into the study and randomly assigned to the control group

$Assessed_{participant}C$ - the number of participants randomly assigned to the control group who completed a 3-month follow-up questionnaire

$Attrition_{participant}T$ - the proportion of participants enrolled in the study and randomly assigned to the treatment group who did not complete a 3-month follow-up questionnaire

$Base_{participant}T$ - the number of participants enrolled into the study and randomly assigned to the treatment group

$Assessed_{participant}T$ - the number of participants randomly assigned to the treatment group who completed a 3-month follow-up questionnaire

Overall and differential attrition will be reported in the findings section of our final impact report.

iii. **Adjustments for multiple comparisons:** Following guidance provided under the grant for our impact analysis plan,⁴² we will adjust for multiple comparisons in all of our primary outcome analyses, regardless of outcome domains. We propose to use the Benjamini-Hochberg method.⁴³ This method controls for the false discovery rate (FDR), which is the expected value of the number of false positive tests divided by the total number of significant tests within a family of tests. The following procedures will be used to implement this adjustment:

1. The p-values generated by our models of the effect of the intervention on our three primary outcome measures will be ranked from smallest to largest, indexed by i (where $i = 1$ for the smallest p-value and $i = k$ for the largest p-value).
2. Beginning with the largest p-value (p_{k1}), we will assess if $p_{k1} < ((i/m)a^*)$, where m = the total number of tests conducted, and a^* = the initial significance value at which we would reject the null hypothesis and the level of false discovery we are willing to accept (in this case, 0.05). The null hypothesis will be rejected and the test will be considered statistically significant if $p_{k1} < ((i/m)a^*)$. If $p_{k1} < ((i/m)a^*)$, all smaller p-values in the list will also be considered statistically significant and the null hypothesis will be rejected for each test. If $p_{k1} \geq ((i/m)a^*)$, the null hypothesis will hold, the test will not be considered statistically significant, and the next largest p-value in the ranked list will be assessed.
3. If the 1st p-value is not statistically significant, the 2nd largest p-value in the list (p_{k2}) will be compared against $(i/m)a^*$. The null hypothesis will be rejected and the test will be considered statistically significant if $p_{k2} < ((i/m)a^*)$. If $p_{k2} < ((i/m)a^*)$, all smaller p-values in the list will also be considered statistically significant and the null hypothesis will be rejected for each test. If $p_{k2} \geq ((i/m)a^*)$, the null hypothesis will hold, the test will not be considered statistically significant, and the next largest p-value in the ranked list will be assessed.
4. If the 2nd p-value is not statistically significant, the 3rd largest p-value in the list (p_{k3}) will be compared against $(i/m)a^*$. The null hypothesis will be rejected and the test will be considered statistically significant if $p_{k3} < ((i/m)a^*)$. If $p_{k3} \geq ((i/m)a^*)$, the null hypothesis will hold and the test will not be considered statistically significant.

iv. **Sensitivity analyses:** We will conduct sensitivity analyses to test the robustness and validity of our benchmark approaches outlined above. These include: (1) excluding covariates; (2) not

⁴² During the January 8, 2019 OAH TPP Tier 2b Group Call on Impact Analysis Plans, the presenters noted that multiple comparison adjustment is required for all model-generated effect estimates of primary outcome measures.

⁴³ This method has been selected because it helps to control the Type 1 error rate without also increasing the Type 2 error rate, which in our view is a serious consideration in preliminary efforts to identify evidence of effectiveness of new approaches. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), 289-300.

imputing or adjusting for missing data; (3) excluding unreliable data; (4) including outliers; and (5) condensing data collection windows to exclude late responders; and (6) using an alternative model specification to estimate program effects.

1. **Without baseline covariates.** Our benchmark approach is to include baseline covariates in our model to improve the precision of our estimates. To test this, we will conduct sensitivity analyses that involve running identical empirical models without the covariates included. Analytic findings for both approaches will be presented alongside each other in an appendix of the impact report.
2. **Without adjusted baseline data.** As outlined in the *Missing data approach* section, our benchmark approach is to adjust baseline data as published guidance suggests that this may produce unbiased impact estimates and maximize the use of available data. We will test this by way of sensitivity analyses that involve running identical empirical models without the adjusted data. Analytic findings for both approaches will be presented alongside each other in an appendix of the impact report.

As outlined in the *Baseline equivalency* section, we will also produce diagnostic estimates of baseline equivalency on the pre-intervention outcome variables according to our benchmark approach and the sensitivity study alongside each other in an appendix of the report.

3. **With unreliable data.** As discussed in the *Data cleaning* section, data for cases that are deemed unreliable are treated as *unit missing* and excluded from benchmark analyses. To test this, we will conduct sensitivity analyses that involve running identical empirical models with the unreliable data included. Analytic findings for both approaches will be presented alongside each other in an appendix of the impact report.
4. **Without outliers.** As discussed in the *Data cleaning* section, extreme data values are investigated and flagged as outliers. Our benchmark analytic approach is to include data flagged as outliers (i.e., extreme values that are not considered invalid) in analysis. We will also conduct sensitivity analyses that exclude these data and report substantive differences in the results section of the report.
5. **Condensed data collection windows.** Our benchmark approach is to include follow-up data from all participants who completed a questionnaire during their open data collection window, regardless of the time point in that window when it was completed. Data collection windows are broad to minimize attrition from the analytic sample. To examine whether or not this influences our results – and, in particular, whether or not study participants who respond later report different outcomes from those who respond earlier – we will conduct an analysis that examines the difference, if any, in response time between treatment and control participants and compares impact estimates for analytic samples without late responders. Late responders will be defined as those participants who complete their three-month questionnaire more than one month after the initiation of the three-month data collection window.
6. **Statistical Modeling.** We have proposed using OLS regression as the benchmark statistical model we intend to use to estimate the program’s effect on the primary outcomes. OLS is robust in large samples to misspecification. OLS is also a conventional approach to modeling dichotomous and count outcomes in evaluation because it produces estimates that are more immediately and readily interpretable, and because it tends to produce results that are substantively identical to the models that technically fit the data better. We will conduct tests that test the validity of this assumption and if there is a

substantive difference in point estimates of interest produced by OLS and logit (or variant of – e.g. Firth logit) or statistical count models (e.g. Poisson and negative binomial families), we will report results of the models that fit the distributional characteristics of the data better (based on diagnostics and log-likelihood statistics).

- g. **Analytic approach for secondary research questions:** We intend to use the same analytic approach as described above to address all secondary research questions, with the exception of our method related to multiple comparisons. Following guidance provided under the grant for our impact analysis plan,⁴⁴ we will not adjust for multiple comparisons in our secondary outcome analyses, regardless of outcome domains.

5) Additional planned analyses

We intend to investigate the following research questions in addition to the primary and secondary research questions described above.

Antecedents of Sexual Behavior

- a. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to participate in *Plan A* (treatment) relative to the offer to participate in *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported perception of risk and severity for pregnancy and HIV/STIs after the end of treatment?
- b. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to participate in *Plan A* (treatment) relative to the offer to participate in *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported intention to use LARC after the end of treatment?
- c. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to participate in *Plan A* (treatment) relative to the offer to participate in *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported intention to use other effective contraception methods (including condoms) after the end of treatment?
- d. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to participate in *Plan A* (treatment) relative to the offer to participate in *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported self-efficacy to communicate with health care providers after the end of treatment?
- e. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to participate in *Plan A* (treatment) relative to the offer to participate in *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported self-efficacy to negotiate condom use after the end of treatment?
- f. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to participate in *Plan A* (treatment) relative to the offer to participate in *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported knowledge or awareness of contraceptive options and LARC after the end of treatment?

Long-term Sexual Behaviors

- a. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported use of long-acting reversible contraceptives (LARC) nine months after receiving the treatment?
- b. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported times having sex without a condom nine months after receiving the treatment?

⁴⁴ See footnote 42.

- c. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported receipt of STI testing nine months after receiving the treatment?
- d. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported use of dual methods of protection (condom and prescription birth control use) during vaginal sex nine months after receiving the treatment?
- e. What is the impact of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (control) on participants' reported use of other effective contraceptive methods (pill, patch, ring, or Depo-Provera) nine months after receiving the treatment?

Effects of Mediators on Primary and Secondary Outcomes of Interest

- a. What are the are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported times having sex without condoms after receiving the treatment considering the following potential mediators:
 - i. Perception of risk and severity for pregnancy and HIV/STIs
 - ii. Intention to use LARC
 - iii. Intention to use other effective contraception methods (including condoms)
 - iv. Provider communication self-efficacy
 - v. Condom negotiation self-efficacy
 - vi. Knowledge or awareness of contraception options and LARC
- b. What are the are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (comparison) on current LARC use after receiving the treatment considering the following potential mediators:
 - i. Perception of risk and severity for pregnancy and HIV/STIs
 - ii. Intention to use LARC
 - iii. Intention to use other effective contraception methods (including condoms)
 - iv. Provider communication self-efficacy
 - v. Condom negotiation self-efficacy
 - vi. Knowledge or awareness of contraception options and LARC
- c. What are the are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (comparison) on STI testing considering the following potential mediators:
 - i. Perception of risk and severity for pregnancy and HIV/STIs
 - ii. Intention to use LARC
 - iii. Intention to use other effective contraception methods (including condoms)
 - iv. Provider communication self-efficacy
 - v. Condom negotiation self-efficacy
 - vi. Knowledge or awareness of contraception options and LARC
- d. What are the are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported use of dual methods of protection considering the following potential mediators:
 - i. Perception of risk and severity for pregnancy and HIV/STIs
 - ii. Intention to use LARC

- iii. Intention to use other effective contraception methods (including condoms)
 - iv. Provider communication self-efficacy
 - v. Condom negotiation self-efficacy
 - vi. Knowledge or awareness of contraception options and LARC
- e. What are the short-term (three-month follow-up) and long-term (nine-month follow-up) impacts of the offer to watch *Plan A* (treatment) relative to the offer to watch *The Toxic Life Cycle of a Cigarette* (comparison) on participants' reported use of other effective contraceptive methods considering the following potential mediators:
 - i. Perception of risk and severity for pregnancy and HIV/STIs
 - ii. Intention to use LARC
 - iii. Intention to use other effective contraception methods (including condoms)
 - iv. Provider communication self-efficacy
 - v. Condom negotiation self-efficacy
 - vi. Knowledge or awareness of contraception options and LARC

Appendix A: Tables

Table 1. Behavioral outcomes used for primary and secondary impact analyses research questions

Outcome name	Description of the outcome, including how it is operationalized	Source of the measure	Timing of measure
Current LARC use	<p>The protective outcome is operationalized as a dichotomous variable indicating whether a person reports currently using LARC (implant or IUD) or not currently using LARC.</p> <p>The measure is calculated from the following item:</p> <ul style="list-style-type: none"> Which of the following methods of prescription birth control are you currently using? <ul style="list-style-type: none"> None: I am not currently using any of these methods Oral contraceptives (for example, the pill) The patch (for example, Ortho Evra) The shot/injection (for example, Depo-Provera) The ring (for example, NuvaRing) The implant (for example, Implanon or Nexplanon) IUD (for example, ParaGard, Skyla, or Mirena) <p>A person who selects either <i>The implant</i> or <i>IUD</i> is given a value of 1 for the measure. A person who selects <i>None</i>, <i>Oral contraceptives</i>, <i>The patch</i>, <i>The shot/injection</i>, or <i>The ring</i> is given a value of 0 for the measure.</p> <p>The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who does not currently use a LARC and 1 indicates a person who does currently use a LARC.</p> <p>Note: The analytic sample will include all respondents who have three-month follow-up data.</p>	<i>Participant Questionnaire</i>	Three-month follow-up (three months after enrollment)
Times having sex without condom	<p>The risk outcome is operationalized as the number of times in the past three months a person reports having any type of sex <u>without</u> using a condom.</p> <p>The measure is calculated from the following items:</p> <ul style="list-style-type: none"> In the past 3 months, how many times have you had vaginal sex without using a condom? 	<i>Participant Questionnaire</i>	Three-month follow-up (three months after enrollment)

- In the past 3 months, how many times have you had oral sex without using a condom?
- In the past 3 months, how many times have you had anal sex without using a condom?

The measure is calculated by summing the total number of times a person reported not using a condom during vaginal, oral and anal sex.

The resulting variable is continuous with values that range from 0 to k , where 0 indicates that a person has not engaged in sex without a condom in the past three months, and k indicates the number of times the person has engaged in sex without a condom (risk behavior) in the past three months.

Note: The analytic sample will include all respondents who have three-month follow-up data. Persons who indicate they have not had sex will be considered to have participated in the risk behavior 0 times (i.e., they did not engage in sex without a condom).

STI testing	<p>The protective outcome is operationalized as a dichotomous variable indicating whether a person reports having been tested for STIs in the past three months or not having been tested.</p> <p>The measure is calculated from the following item:</p> <ul style="list-style-type: none"> • Have you been tested for STIs/STDs other than HIV in the past 3 months? <p>A person who selects either <i>Yes</i> is given a value of 1 for the measure. A person who selects <i>No</i> is given a value of 0 for the measure.</p> <p>The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who has not been tested for STIs/STDs in the past three months and 1 indicates a person who has been tested for STIs/STDs.</p> <p>Note: The analytic sample will include all respondents who have three-month follow-up data.</p>	<i>Participant Questionnaire</i>	Three-month follow-up (three months after enrollment)
Use of dual methods of protection	<p>The protective outcome is operationalized as the proportion of times in the past three months a person reports using both a condom and prescription birth control during vaginal sex out of total times reported having vaginal sex.</p>	<i>Participant Questionnaire</i>	Three-month follow-up (three months after enrollment)

The measure is calculated from the following two items:

- In the past three months, how many times have you had vaginal sex?
- Please think about how many times you said you had vaginal sex in question 20. Of the times you had vaginal sex, how many times would you say you used dual methods of protection – that is, how many times did you use both a condom and one of the listed methods of prescription birth control at the same time?

The resulting variable is a continuous proportion with values that range from 0 to 1, where 0 indicates that a person has used dual methods of protection 0% of the time during vaginal sex in the past three months, and 1 indicates the person has used dual methods of protection 100% of the time during vaginal sex in the past three months.

Note: The analytic sample will only include respondents who have three-month follow-up data and who indicate in their three-month follow-up questionnaire that they have had vaginal sex in the past three months. Persons who indicate they have never had vaginal sex or have not had vaginal sex in the past three months will be excluded from the construction of this measure.

Use of other effective contraceptive methods

The protective outcome is operationalized as a dichotomous variable indicating whether a person reports currently using prescription birth control methods beyond LARC or not currently using any prescription birth control methods.

*Participant
Questionnaire*

Three-month follow-up
(three months after enrollment)

The measure is calculated from the following item:

- Which of the following methods of prescription birth control are you currently using?
 - None: I am not currently using any of these methods
 - Oral contraceptives (for example, the pill)
 - The patch (for example, Ortho Evra)
 - The shot/injection (for example, Depo-Provera)
 - The ring (for example, NuvaRing)
 - The implant (for example, Implanon or Nexplanon)
 - IUD (for example, ParaGard, Skyla, or Mirena)

A person who selects either *Oral contraceptives*, *The patch*, *The shot/injection*, or *The ring* is given a value of 1 for the measure. A person who selects *None* is given a value of 0 for the measure.

The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who does not currently use any prescription birth control methods and 1 indicates a person who uses non-LARC prescription birth control methods.

Note: The analytic sample will only include respondents who have three-month follow-up data and who indicate in their three-month follow-up questionnaire that they are currently not using any birth control or using one of the non-LARC methods. Persons who indicate they use LARC methods will be excluded from the construction of this measure.

Table 2. Covariates included in primary and secondary impact analyses

Covariate	Description of the covariate and how it will be used as a covariate in the analysis	Rationale for inclusion
<i>Behavioral outcomes at baseline</i>		
Current LARC use	<p>The protective outcome is operationalized as a dichotomous variable indicating whether a person reports currently using LARC (implant or IUD) or not currently using LARC.</p> <p>The measure is calculated from the following item:</p> <ul style="list-style-type: none"> Which of the following methods of prescription birth control are you currently using? <ul style="list-style-type: none"> None: I am not currently using any of these methods Oral contraceptives (for example, the pill) The patch (for example, Ortho Evra) The shot/injection (for example, Depo-Provera) The ring (for example, NuvaRing) The implant (for example, Implanon or Nexplanon) IUD (for example, ParaGard, Skyla, or Mirena) <p>A person who selects either <i>The implant</i> or <i>IUD</i> is given a value of 1 for the measure. A person who selects <i>None</i>, <i>Oral contraceptives</i>, <i>The patch</i>, <i>The shot/injection</i>, or <i>The ring</i> is given a value of 0 for the measure.</p> <p>The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who does not currently use a LARC and 1 indicates a person who does currently use a LARC.</p>	Current LARC use is included in the primary impact analysis as the pre-intervention or baseline measure of the behavioral outcome; it is included in the models so that individual-level change or difference can be assessed at the three-month follow-up.
Times having sex without condoms	<p>The risk outcome is operationalized as the number of times in the past three months a person reports having any type of sex <u>without</u> using a condom.</p> <p>The measure is calculated from the following items:</p> <ul style="list-style-type: none"> In the past 3 months, how many times have you had vaginal sex without using a condom? In the past 3 months, how many times have you had oral sex without using a condom? 	Times having sex without condoms is included in the primary impact analysis as the pre-intervention or baseline measure of the behavioral outcome; it is included in the models so that individual-level change or difference can be assessed at the three-month follow-up.

Covariate	Description of the covariate and how it will be used as a covariate in the analysis	Rationale for inclusion
	<ul style="list-style-type: none"> In the past 3 months, how many times have you had anal sex without using a condom? <p>The measure is calculated by summing the total number of times a person reported not using a condom during vaginal, oral and anal sex.</p> <p>The resulting variable is continuous with values that range from 0 to k, where 0 indicates that a person has not engaged in sex without a condom in the past three months, and k indicates the number of times the person has engaged in sex without a condom (risk behavior) in the past three months.</p>	
STI testing	<p>The protective outcome is operationalized as a dichotomous variable indicating whether a person reports having been tested for STIs in the past three months or not having been tested.</p> <p>The measure is calculated from the following item:</p> <ul style="list-style-type: none"> Have you been tested for STIs/STDs other than HIV in the past 3 months? <p>A person who selects either <i>Yes</i> is given a value of 1 for the measure. A person who selects <i>No</i> is given a value of 0 for the measure.</p> <p>The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who has not been tested for STIs/STDs in the past three months and 1 indicates a person who has been tested for STIs/STDs.</p>	<p>STI testing is included in the primary impact analysis as the pre-intervention or baseline measure of the behavioral outcome; it is included in the models so that individual-level change or difference can be assessed at the three-month follow-up.</p>
Use of dual methods of protection	<p>The protective outcome is operationalized as the proportion of times in the past three months a person reports using both a condom and prescription birth control during vaginal sex out of total times reported having vaginal sex.</p> <p>The measure is calculated from the following two items:</p> <ul style="list-style-type: none"> In the past three months, how many times have you had vaginal sex? Please think about how many times you said you had vaginal sex in question 20. Of the times you had vaginal sex, how many 	<p>Use of dual methods of protection is included in the secondary impact analysis as the pre-intervention or baseline measure of the behavioral outcome; it is included in the models so that individual-level change or difference can be assessed at the three-month follow-up.</p>

Covariate	Description of the covariate and how it will be used as a covariate in the analysis	Rationale for inclusion
	<p>times would you say you used dual methods of protection – that is, how many times did you use both a condom and one of the listed methods of prescription birth control at the same time?</p> <p>The resulting variable is a continuous proportion with values that range from 0 to 1, where 0 indicates that a person has used dual methods of protection 0% of the time during vaginal sex in the past three months, and 1 indicates the person has used dual methods of protection 100% of the time during vaginal sex in the past three months.</p>	
Use of other effective contraceptive methods	<p>The protective outcome is operationalized as a dichotomous variable indicating whether a person reports currently using prescription birth control methods beyond LARC or not currently using any prescription birth control methods.</p> <p>The measure is calculated from the following item:</p> <ul style="list-style-type: none"> • Which of the following methods of prescription birth control are you currently using? <ul style="list-style-type: none"> ○ None: I am not currently using any of these methods ○ Oral contraceptives (for example, the pill) ○ The patch (for example, Ortho Evra) ○ The shot/injection (for example, Depo-Provera) ○ The ring (for example, NuvaRing) ○ The implant (for example, Implanon or Nexplanon) ○ IUD (for example, ParaGard, Skyla, or Mirena) <p>A person who selects either <i>Oral contraceptives</i>, <i>The patch</i>, <i>The shot/injection</i>, or <i>The ring</i> is given a value of 1 for the measure. A person who selects <i>None</i> is given a value of 0 for the measure.</p> <p>The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who does not currently use any prescription birth control methods and 1 indicates a person who uses non-LARC prescription birth control methods.</p>	Use of other effective contraceptive methods is included in the secondary impact analysis as the pre-intervention or baseline measure of the behavioral outcome; it is included in the models so that individual-level change or difference can be assessed at the three-month follow-up.
Individual level covariates		

Covariate	Description of the covariate and how it will be used as a covariate in the analysis	Rationale for inclusion
Age	<p>The variable is measured as the respondent's age in years at screening.</p> <p>The measure is constructed from the following item on the <i>VHS Eligibility Screening Form</i>:</p> <ul style="list-style-type: none"> • What is your date of birth? <p>The variable is calculated by subtracting the reported date of birth given from the date when the screening was completed.</p> <p>The resulting variable is continuous with values ranging from 18 to 19.</p>	<p>Research has shown that likelihood of engaging in sex increases with age and use of condoms declines (Brewster 1999; Kirby 2007; Miller et al 1998; Scott-Jones and White 1990)</p>
Black/African American	<p>The measure is operationalized as a dummy variable, where 0 = identify as another race/do not identify race; 1 = identify as Black/African American.</p> <p>The measure is taken directly from the following item on the <i>VHS Eligibility Screening Form</i>:</p> <ul style="list-style-type: none"> • What is your race and ethnicity? (Study coordinators can select more than one based on participant's response) <ul style="list-style-type: none"> ○ White ○ Black or African American ○ Hispanic, Latino, or of Spanish origin ○ American Indian or Alaska Native ○ Asian ○ Native Hawaiian or Pacific Islander ○ Unknown ○ Some other race/ethnicity (specify) <p>Variable will be coded as 1 if participant self-identified as Black or African American, regardless as to whether other races/ethnicities are specified; if Black or African American is not selected, the response will be coded as 0.</p>	<p>Research has shown that Black/African American and Hispanic adolescents are more likely to engage in sex during adolescence, initiate sexual activity at a younger age, and less likely to use contraception of any kind (Blum 2000; Brewster 1999; Hogan et al 2000; Kirby 2007; Scott Jones and White 1990)</p>

Covariate	Description of the covariate and how it will be used as a covariate in the analysis	Rationale for inclusion
Hispanic, Latino or of Spanish origin	<p>The measure is operationalized as a dummy variable, where respondents who identify as Hispanic, Latino, or of Spanish origin are coded as 1 and coded as 0 otherwise.</p> <p>The measure is taken directly from the following item on the <i>VHS Eligibility Screening Form</i>:</p> <ul style="list-style-type: none"> • What is your race and ethnicity? (Study coordinators can select more than one option based on participant's response) <ul style="list-style-type: none"> ○ White ○ Black or African American ○ Hispanic, Latino, or of Spanish origin ○ American Indian or Alaska Native ○ Asian ○ Native Hawaiian or Pacific Islander ○ Unknown ○ Some other race/ethnicity (specify) <p>Variable will be coded as 1 if participant self-identified as Hispanic, Latino, or of Spanish origin, regardless as to whether other races/ethnicities are specified; if Hispanic, Latino, or of Spanish origin is not selected, the response will be coded as 0.</p>	
High school education	<p>This measure is operationalized as a dummy variable, where individuals with a high school education are coded as 1 and all others are coded as 0.</p> <p>The measure is taken directly from the following item on the <i>Baseline Participant Questionnaire</i>:</p> <ul style="list-style-type: none"> • What is the highest degree or level of school you have completed (if currently enrolled, select the previous grade completed or degree received)? <ul style="list-style-type: none"> ○ Grade 9-12 (please specify grade) ○ GED or alternative credential ○ Some college credit, but less than 1 year of credit 	

Covariate	Description of the covariate and how it will be used as a covariate in the analysis	Rationale for inclusion
	<ul style="list-style-type: none"> ○ 1 or more years of college credit, no degree ○ Associate's degree ○ Bachelor's degree ○ Other (specify) <p>Participants who select <i>Grade 9-12</i> will be coded as 0=has not completed high school; all others will be coded as 1=has completed high school.</p>	
Blocking Covariates		
Region	<p>The measure is operationalized a set of $n-1$ dummy variables, where n refers to the number of VHS regions over the course of the evaluation period.</p> <p>Data for the measure are obtained from the <i>Enrollment Log Database</i>.</p> <p>Each dummy will be coded as 1 if the individual is enrolled in a particular region and 0 otherwise. Region 1 is the reference variable. Dummy variables will be grand mean centered so that the intercept will then reflect the un-weighted mean site effect.</p>	Randomization is blocked by region.

Table 3: Data Editing Rules

The following table provides PRG's general rules for editing data based upon responses given.

Category	Data editing rule
No response given to an item (coded as .f)	If data from a related variable can be used to infer a value, data will be logically edited. Otherwise, the value will be left as missing.
Invalid items (coded as .i)	Adjust missing baseline values
Outlying items (Outlier indicator variable coded as 1)	Keep in benchmark analysis; run sensitivity analyses excluding outliers
Inconsistent across-time items (coded as .k)	Adjust missing baseline values
Unreliable cases (Unreliable indicator variable codes as 1)	Exclude case from benchmark analysis; run sensitivity analyses including unreliable cases

Appendix B: Evaluation Abstract Submitted to OAH in September 2017

EVALUATION ABSTRACT : THE EVALUATION OF 'PLAN A' IN CALIFORNIA

Grantee

Grantee Name: The Policy & Research Group
Project Lead: Lynne Jenner
Email address: ljenner@policyandresearch.com

Evaluator

Evaluator's Organization: The Policy & Research Group
Evaluator Lead: Eric Jenner
Email address: ejenner@policyandresearch.com

Intervention Name

"Plan A"

Intervention Description

"Plan A" is a 23-minute video intervention designed for 18- to 19-year-old African American and Hispanic or Latina women that promotes effective contraceptive use, condom use for dual contraception and human immunodeficiency virus/sexually transmitted disease (HIV/STD) prevention, and HIV/STD testing. The video aims to develop sexual health intentions, knowledge, and self-efficacy for communicating with providers about different contraceptive options that have been proven effective, such as long-acting reversible contraception (LARC).

The video is delivered on laptops or personal electronic devices in a private room or area of a reproductive health clinic. The intervention is designed to have maximum impact when viewed just before a reproductive health visit. The video developers believe that the waiting time before a clinic visit is a moment when the target group will be most receptive to the informational and motivational messages of the intervention. "Plan A" intends to improve sexual health outcomes by empowering viewers to understand their options and communicate their needs to their health providers to get the most out of their experience at a reproductive health clinic.

Comparison Condition

"The Toxic Life Cycle of a Cigarette"

Comparison Condition Description

"The Toxic Life Cycle of a Cigarette" is a 17-minute video that details the negative effects that cigarettes have on the environment and on people who manufacture and use cigarettes. The informational video uses both narration and interviews to educate viewers on the dangers that cigarettes pose. The video is delivered on laptops or personal electronic devices in a private room or area of a Planned Parenthood clinic. The video includes no sexual or reproductive health content.

Behavioral Outcomes

Use of LARC (intrauterine device or implant), use of condoms, STD testing, use of dual methods of protection (condom and prescription birth control), use of other effective contraception methods (consistent and correct use of pill, patch, ring, or Depo-Provera)

Non-behavioral Outcomes

Perception of risk and severity for pregnancy and HIV/STD infection, intention to use LARC, intention to use other effective contraception methods (including condoms), provider communication self-efficacy (or positive outcome expectation from provider interaction), condom negotiation self-efficacy, knowledge or awareness of contraception options and LARC

Sample and Setting

The study is being conducted in ten Planned Parenthood clinics in five different regions of California—Fresno, Oakland, Merced, Bakersfield, and Sacramento. All of the clinics are part of the Planned Parenthood Mar Monte organization. A study coordinator will screen for eligibility all women with reproductive health appointments (including walk-in appointments) at study clinics. Study coordinators are nonclinical employees of Planned Parenthood Mar Monte. None of the study coordinators know the specific reason youth are visiting the clinic, though they might know the type of appointment the participant has scheduled. The intent-to-treat sample will be comprised of eligible teens who are enrolled into the study during the two-year implementation period. To be eligible, participants must (1) be female; (2) be 18 or 19 years old; (3) self-identify as either Hispanic, Latina, or African American; (4) be visiting a reproductive health clinician or provider; (5) be deemed appropriate for the study by clinic staff with regards to physical and mental health capacity; (6) consent to participate in the study; (7) not knowingly be pregnant; and (8) not be trying to get pregnant. The target sample size for the study is 1,770 women.

Research Design and Data Collection

The study is an individual randomized controlled trial in which eligible, consenting participants are randomly assigned to either the intervention (“Plan A”) or comparison (“The Toxic Life Cycle of a Cigarette”) groups. Random assignment is conducted following consent and study enrollment and before the administration of the baseline questionnaire. The evaluators are responsible for coordinating and verifying random assignment. The standard approach is to carry out random assignment electronically through a web-based survey platform; however, the contingency plan is for study coordinators to use assignment envelopes if Internet access is not available or insufficient at an administration site for any reason.

Participants in both the intervention and comparison groups will receive a baseline survey (before the intervention), 3-month follow-up (post-baseline), and 9-month follow-up (post-baseline). Data collection procedures will be the same for both treatment and comparison groups. In-person data collection will be the preferred mode for all data points. In these instances, participants will complete questionnaires (available in English and Spanish) electronically through a web-based survey form on a computer in a private room or space; however, paper questionnaires will be available when needed or preferred by participants. When participants cannot meet in person to complete follow-up data collection, participants will be offered alternative methods to complete the questionnaires (for example, online using a personal electronic device or by telephone). Data collection windows for both follow-up questionnaires will be four months. The evaluators will track when participants respond and sensitivity analyses will examine whether outcomes differ for participants who responded on time (within the first month of the data collection window) and late (in the second, third, or fourth months of the window).

For the implementation evaluation, the evaluators will collect data on fidelity (participant-reported understanding and participation or engagement), attendance (percentage of video watched), and quality (participant-reported overall quality). The web-based video platform (Wistia) will automatically log attendance and the Plan A Feedback Questionnaire will collect fidelity and quality data; a randomly selected 10 percent of study participants in the treatment group will complete that questionnaire electronically through a web-based survey form on a computer directly after viewing the “Plan A” video.

Schedule/Timeline

Sample enrollment and baseline data collection began in June 2016 and will end in October 2018. The 3-month follow-up data collection began in September 2016 and will end in May 2019. The 9-month follow-up data collection will begin in March 2017 and will end in November 2019.

Prepared for the Office of Adolescent Health by Mathematica Policy Research under contract #GS-10F-0050L, Task Order No. HHSP233201300416G with Policy and Research, LLC with grant #6 TP2AH000036.