



Teen Pregnancy Prevention Replication Study: Short-Term Impacts of *iCuídate!*

| EVALUATION REPORT |

OCTOBER 2016



***¡Cúídate!:* Interim**

Impact Report

**Teen Pregnancy
Prevention
Replication Study**

October 2016

Prepared for:

Amy Farb
The Office of Adolescent
Health

Lisa Trivits

Sarah Oberlander

The Office of the Assistant
Secretary for Planning and
Evaluation

Submitted by:

Meredith Kelsey (Abt)

*Jean Layzer (Belmont
Research Associates)*

Cristofer Price (Abt)

Randall Juras (Abt)

Abt Associates

55 Wheeler St.

Cambridge, MA 02138

In Association with:

**Decision Information
Resources, Inc.**

and

Belmont Research Associates

Recommended Citation

Kelsey, M., Layzer, J., Price, C., and Juras, R. “¡Cúdate! Short-Term Impact Report: Findings from the Teen Pregnancy Prevention Replication Study”. Cambridge, MA: Abt Associates Inc., October 2016. This document is available at <https://aspe.hhs.gov/teen-pregnancy-prevention-tpp-replication-study>

This document is available at <https://aspe.hhs.gov/teen-pregnancy-prevention-tpp-replication-study>

This document was prepared under contract number HHSP23320095624WC Order No. HHSP23337011T (awarded in September 2011) from the U.S. Department of Health and Human Services, Office of Adolescent Health and the Office of the Assistant Secretary for Planning and Evaluation. Any statements expressed are those of the authors and do not necessarily reflect the views of the Office of Adolescent Health, the Office of the Assistant Secretary for Planning and Evaluation, or the U.S. Department of Health and Human Services.

Acknowledgements

We would like to thank many individuals and organizations who helped to make this evaluation possible. First, we thank the staff from the Office of Adolescent Health (OAH) within the U.S. Department of Health and Human Services (HHS) for their support throughout the project. In particular Amy Farb, the OAH project officer for the Teen Pregnancy Prevention Replication Study, provided critical guidance throughout the project. We also thank Amy Margolis, the many program officers and Alexandra Warner. Lisa Trivits and Sarah Oberlander, the project officers for the TPP Replication Study from the Office of the Assistant Secretary for Planning and Evaluation (ASPE) provided critical direction and support throughout the project and helped to ensure the success of the study.

We would like to thank our project partners Decision Information Resources (DIR) for their commitment and hard work in support of the study and data collection efforts, along with Innovative Clinical Research Solutions (ICRS). Belmont Research Associates provided invaluable leadership and project direction, and helped navigate many obstacles and establish critical relationships on the ground. We would like to acknowledge each of the grantee sites whose close partnership and commitment were essential to the success of the study. Project staff not only implemented the program, but participated in interviews and remained dedicated to the study throughout.

At Abt, the project was supported by a team of individuals who worked closely with each of the grantee sites. Jackie Mendez and Carolyn Layzer provided essential support for implementation of the study. Rob Olsen provided guidance and support during the initial study design phase, and Fatih Unlu provided valuable feedback on the report. Thanks to Dan Weiss and Heather He for programming and technical assistance, and Stefanie Falzone and Jan Nicholson for producing the report.

And finally, a tremendous thanks to all of the youth who participated in the study. Without them, this entire endeavor would not have been possible.

Contents

1. Introduction..... 1

1.1 The Replication Study..... 1

1.2 The Three Models Replicated..... 2

1.3 Focus of This Report..... 2

2. The Program Model: ¡Cuidate! 3

2.1 ¡Cuidate! Logic Model 4

2.2 Evidence of Effectiveness 5

3. Evaluation Design 7

3.1 Research Questions 7

3.2 Study Design 7

3.2.1 Selection of Replication Grantees..... 7

3.2.2 Site-Specific Program Designs 8

3.2.3 Settings for the Program 9

3.2.4 Recruitment and Random Assignment..... 9

3.2.5 Treatment and Control Conditions..... 11

3.3 Data Collection Strategies and Measures..... 12

3.3.1 Data Collection Strategy 12

3.3.2 Measures 13

3.4 Analytic Approach 16

3.4.1 Estimation of Impacts for the Full Sample 16

3.4.2 Site-level Analyses..... 18

3.4.3 Subgroup Analyses 18

3.4.4 Handling Missing Data 18

3.4.5 Addressing Multiple Comparisons..... 19

4. Results..... 21

4.1 Program Implementation..... 21

4.1.1 Staff Hiring and Training 21

4.1.2 Implementing the Program with Fidelity 22

4.1.3 Participant Attendance and Engagement 22

4.2 Sample Characteristics..... 23

4.2.1 Study Sample 23

4.2.2 Comparability of the Treatment and Control Groups at Baseline..... 26

4.3 Program Impacts on Exposure to Sexual Health Information..... 26

4.4 Intermediate Outcomes 28

4.4.1 Knowledge 28

4.4.2 Attitudes 30

4.4.3 Motivation..... 32

4.4.4 Intentions..... 33

4.4.5 Skills 34

4.5 Youth Sexual Behavior and Sexual Risk 34

5. Discussion 37

CONTENTS

References.....	40
Appendix A: Site-Level Impacts.....	42
Appendix B: Subgroup Impacts	51
Appendix C: Measures	55
Appendix D: Supporting Tables	65

1. Introduction

Reducing rates of unplanned teen pregnancy and of sexually-transmitted infections (STIs) are priorities for the U.S. Department of Health and Human Services (DHHS). To achieve this goal, the Department is investing in evidence-based pregnancy reduction strategies and targeting populations at highest risk for teen pregnancy. The federal Teen Pregnancy Prevention (TPP) Program, administered by the Office of Adolescent Health (OAH), includes funding for programs that are intended to address high rates of teenage pregnancy by (1) replicating evidence-based models and (2) testing innovative strategies.

The TPP Program was authorized in 2010 as part of the larger Teen Pregnancy Prevention Initiative. The program initially included \$100 million in annual funding to support programming. Of these funds, \$75 million were available annually to support five-year grants for replicating 28 program models that prior rigorous evaluations had shown to be effective. These program models were identified through a systematic, comprehensive review of the literature on teen pregnancy, STIs, and sexual risk behaviors (Kappeler & Farb, 2014).

The TPP program acknowledges the limitations of existing research and the need for additional research on programs, citing lessons learned from a comprehensive evidence review such as an absence of independent evaluations and a limited number of program replications (Goesling et al., 2014). In the review, the evidence for many of the programs included on the list rests on a single study of effectiveness, often conducted a long time ago and with a single population. A program may work in one location with a particular population, but that does not necessarily mean it will be effective in another. Further, implementing a program with fidelity often competes with the need to adapt to local conditions on the ground. A carefully designed study of multiple replications of different program models is an important contribution to the existing research.

1.1 The Replication Study

The TPP Replication Study is being conducted for OAH, under contract with the Office of the Assistant Secretary for Planning and Evaluation (ASPE) by Abt Associates and its subcontractors, Belmont Research Associates, Decision Information Resources (DIR), and CiviCore.¹ The study has two major components: an Impact Study and an Implementation Study.

Impact Study. Through a series of rigorous experimental design evaluations, the study tests multiple replications of several widely-used evidence-based program models to determine their effectiveness across different settings and populations. The strategy of selecting multiple replications of each program model allows us to pool data across the three replication sites to assess impacts on behavioral outcomes such as pregnancy and for key subgroups (e.g., those based on age and sexual experience). In addition, the strategy lets us examine variation in impacts across replications for each program model and gather evidence about the generalizability of findings on program effectiveness.

¹ The study is also referred to as The Teen Health Empowerment Study in the field with program staff and study participants.

Implementation Study. A comprehensive implementation study will provide information about the contexts in which the evidence-based programs were implemented, the challenges faced in implementing them, and the aspects of program implementation that are associated with program impacts.

1.2 The Three Models Replicated

OAH, with its ASPE partners, selected three program models from the first round of TPP-funded grants to test and replicate: the *Safer Sex Intervention* (a clinic-based HIV/STI prevention program for high-risk adolescent females); *Reducing the Risk* (a sexual health education curriculum); and *¡Cuidate!* (an HIV/STI risk reduction program targeting Latino youth). These programs were selected based on the breadth and scale of the proposed replication effort. All three were proposed for replication by at least five grantees.² In addition, the three represent a range of targeting and service strategies, as well as some variation in the settings in which services are provided.

1.3 Focus of This Report

This report, which focuses on *¡Cuidate!*, is one in a series of reports that present findings on the implementation and effectiveness of the three program models. The report presents findings from the first of two follow-up surveys designed to examine the short-term and longer-term impacts of *¡Cuidate!*. Two companion reports will examine the short-term impacts (six to 12 months post-baseline³) of the remaining program models in the study. Three final impact reports will present findings on the longer-term impacts (18 to 24 months post-baseline) of all three program models. Three Implementation Study reports will document the implementation of each of the program models. In addition, nine site profiles provide an overview of program implementation as well as descriptive information about the study participants at baseline in each site.⁴

² Of the 28 program models in the TPP Program, the *Teen Outreach Program (TOP)* is the most frequently replicated. Seven independent evaluations of TOP were conducted as a condition of those grants. For this reason, it was excluded from consideration for the TPP Replication Study. *Becoming a Responsible Teen (BART)*, another widely used model, was also excluded because it had already undergone several evaluations.

³ Where “baseline” means the point at which each study participant entered the study

⁴ The profiles are available at <https://aspe.hhs.gov/basic-report/tpp-replication-study>.

2. The Program Model: ¡Cuidate!

¡Cuidate! is an HIV/AIDS prevention curriculum adapted from the *Be Proud! Be Responsible!* curriculum and culturally tailored for Hispanic youth. It aims to reduce the risk of sexually transmitted infections (STIs), in particular HIV, by affecting sexual behaviors such as sexual intercourse, number of sexual partners, and condom use. Six 60-minute modules are delivered in small groups of 6–10 youth, led by a trained adult facilitator who is bilingual in English and Spanish, although the program is delivered in English only.

¡Cuidate! was originally tested in an after-school setting on consecutive weekends, but can be delivered in other settings, and on different schedules (Villaruel et al, 2005). The curriculum modules are delivered in participatory, interactive sessions. Each session weaves in the theme of *Taking Care*—of oneself, one’s partner, family and community. Exhibit 2.1 shows the topics covered in each of the six modules and links them to the program’s core elements.

Exhibit 2.1: ¡Cuidate! Modules, Topics and Core Element(s) Covered

Module	Topic/Activities	Core Elements Addressed ^a
Introduction and Overview	<ul style="list-style-type: none"> • Getting to Know You • Talking Circle • Creating group rules • HIV/AIDS • What it means to be Latino/Latina • Cultural values • What Latinos think about HIV/AIDS and safer sex 	1, 2, 3, 6
Building HIV Knowledge	<ul style="list-style-type: none"> • View ¡Cuidate! video • Myths and facts 	1, 2, 3, 6
Understanding Vulnerability to HIV Infection	<ul style="list-style-type: none"> • Acknowledging the threat of HIV/AIDS • Latino cultural values and HIV • “A Romance” (role play) • La Lotería • Talking Circle 	1, 2, 3, 5, 6
Attitudes and Beliefs about HIV/AIDS	<ul style="list-style-type: none"> • Welcome and Talking Circle • Music and discussion • Quién es más macho? Quién es más mujer? • Adolescent vulnerability to HIV • La Zona Peligrosa 	1, 2, 3, 6
Building Condom-Skills	<ul style="list-style-type: none"> • Discussing condoms • Condom-use skills • Overcoming barriers to condom use • What gets in the way of caring behavior? • Condom line-up 	1, 3, 4, 5, 6
Building Negotiation and Refusal Skills	<ul style="list-style-type: none"> • No Hay Razon • How to use the S.W.A.T. technique and scripted role plays • AIDS Jeopardy game • Talking Circle 	1, 2, 3, 5, 6

Source: ¡Cuidate! Starter Kit

^a Core elements addressed are as follows.

- 1) Incorporating the theme of taking care of oneself and one’s partner, family and community throughout the program.
- 2) Using culturally- and linguistically-appropriate materials and activities to show and emphasize core Latino cultural values, specifically *familialism* and gender roles and how those are consistent with safer sex behavior.

- 3) Incorporating activities that increase knowledge and influence positive attitudes, beliefs and self-efficacy regarding HIV sexual risk-reduction behaviors.
- 4) Modeling and practicing the effective use of condoms.
- 5) Building participants' skills in problem-solving, negotiation of safe sex, and refusal of unsafe sex.
- 6) Delivering sessions in highly participatory, interactive small groups.

2.1 ¡Cuidate! Logic Model

The materials used in the sessions emphasize core Hispanic values and feelings, and link them to safer sexual behavior. The facilitator demonstrates correct condom use, and teaches negotiation and refusal skills. Youth are exposed to information about HIV/STI transmission and prevention. Through active participation in discussions, sharing ideas and feelings and role-playing situations in which they may be pressured to have unwanted or unsafe sex, participants are exposed to more information to increase their knowledge and understanding of how to avoid unsafe sexual behaviors, improve their attitudes, values, and beliefs, and increase their motivation to delay childbearing.

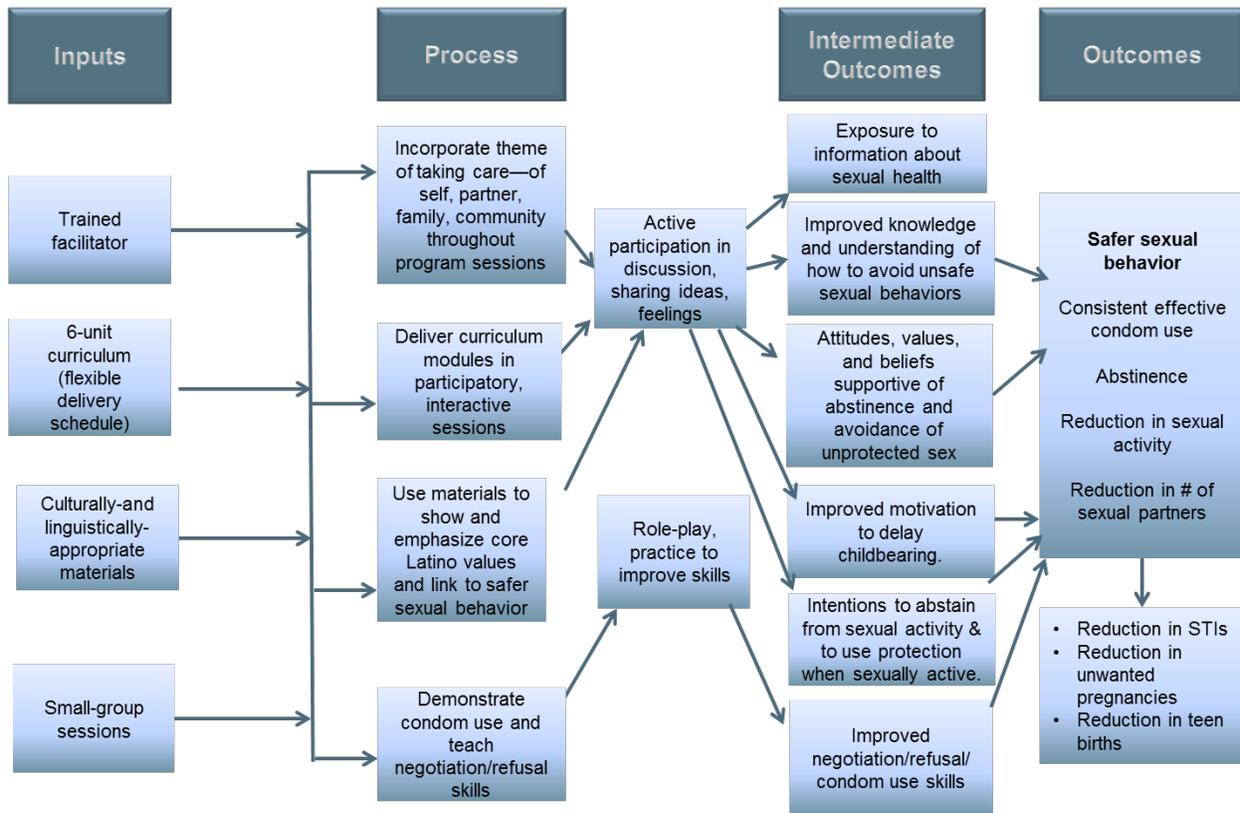
These discussions, together with repeated role-play activities, strengthen intentions to abstain from sexual activity and to use protection. Role plays also support the acquisition of skills youth need to deal with unwanted pressures and risky situations, refuse unsafe sex and negotiate safer sex, and use condoms correctly. These changes in intermediate outcomes (exposure to information, positive changes in knowledge, attitudes, motivation, intentions, and skills) are hypothesized to lead to the safer behavioral outcomes that the program seeks to achieve: correct and consistent use of condoms, abstinence from sex, reduced sexual activity, and reduced number of partners. Safer sexual behavior is ultimately expected to result in reduction in the rates of STIs among teens as well as reduction in pregnancy rates and births.⁵

Exhibit 2.2 shows the program elements, the hypothesized outcomes, and the pathways by which the program seeks to achieve these outcomes.

In Section 3.2, we describe in more detail the modifications to the program model proposed by each of the three organizations implementing the program replications. Modifications were made to comply with state mandates, to address gaps in program content, or to accommodate local constraints, while adhering to the core elements of the model.

⁵ In 2012, the curriculum was revised to include material on pregnancy prevention. However, the grantees whose projects were funded in 2010 were trained on and implemented the original curriculum, supplementing it with additional sessions on pregnancy prevention or weaving that theme into existing sessions.

Exhibit 2.2: Logic Model for ¡Cuidate!



2.2 Evidence of Effectiveness

¡Cuidate! is one of a handful of programs that address the issue of sexual risk behavior in Hispanic adolescents, and one of the few that met the standards for inclusion on the list of evidence-based programs from which TPP grantees could choose (DHHS, 2010). Aside from the single research study of ¡Cuidate! cited in the evidence review (Villaruel et al, 2006), few studies of pregnancy or HIV/STI prevention have focused on Hispanic youth, although a few have included substantial numbers of Hispanic adolescents as part of a larger population. ¡Cuidate! was developed more than a decade ago in response to several concerns.

Hispanics, and Hispanic youth in particular, constitute a large growing portion of the US population. Since 1995, Hispanic children have been the largest minority group of children in the US. By 2011, 23 percent of the population under the age of 18 was Hispanic (US Census Bureau, 2012). This population is disproportionately affected by HIV/AIDS. In 2010, the rate of HIV/AIDS in Hispanic adults/adolescents was more than 3 times as high as in the corresponding white population and Hispanics accounted for 21 percent of all new HIV infections (CDC, 2012). This disproportion in rates has remained stable since 2001. Among possible reasons for the disparity are: Hispanic adolescents are less likely to use condoms and more likely to report multiple sexual partners (CDC-P, 2004). Lower rates of condom use by Hispanic youth continue to be a problem; in 2011 only 54.9 percent of Hispanic students reported using a condom during sex, compared with 63.3 percent and 62.4 percent of White and African-American students respectively (American Academy of Pediatrics, 2013).

¡Cuidate! is based on *Be Proud, Be Responsible!*, an HIV reduction program that was shown to be effective for African-American youth and which is also on the list of evidence-based programs (Jemmott, J. et al 1999; Jemmott, J. 1992; Goesling et al, 2014). It was tested by the developer among mostly Puerto Rican youth in Philadelphia. Youth in the program reported fewer incidents of sexual intercourse, fewer sex partners, and fewer days of unprotected intercourse than youth assigned to a health promotion program (Villaruel et al. 2006). In an earlier paper (Villaruel et al., 2005), the developer notes the diversity among Hispanics in the US and its implications for adolescent risk behavior (HIV/AIDS rates are highest among Puerto Ricans in the Northeast and lowest among Mexican-Americans in the West), but does not restrict the generalizability of the study's findings to the Hispanic population on which it was tested. Since two of the three replications in the TPP Replication Study were conducted in Western states, with primarily Mexican-American populations, this evaluation sheds some light on the generalizability of the intervention.

3. Evaluation Design

The impact study is designed to estimate the effects of three replications of *¡Cuídate!*⁶ It addresses questions about the replications' effects on sexual risk behaviors as well as on the intermediate outcomes the logic model predicts will lead to the behavioral outcomes that *¡Cuídate!* seeks to achieve. This report is focused on program effects six months after study enrollment and is guided by the research questions below.⁷

3.1 Research Questions

- *Did ¡Cuídate! increase student exposure to information on reproductive health, contraceptive methods, and STI transmission and prevention?*
- *Did ¡Cuídate! improve teens' knowledge and understanding of reproductive health, risky sexual behavior, pregnancy prevention, and the transmission and prevention of sexually transmitted infections (STIs)?*
- *Did ¡Cuídate! have positive effects on teens' attitudes towards sexual activity, birth control and condom use?*
- *Did ¡Cuídate! increase motivation to delay childbearing?*
- *Did ¡Cuídate! increase intentions to avoid risky sexual behavior?*
- *Did ¡Cuídate! increase teens' confidence in their ability to refuse unwanted sex and to negotiate safe sex?*
- *Did ¡Cuídate! lead teens to delay sexual initiation and reduce risky sexual activity?*

3.2 Study Design

In each of the replication sites, the study employed an experimental design in which students were randomly assigned to a group that received the *¡Cuídate!* intervention or to a control group that did not. This section describes the selection of the three replication sites, site-specific program designs, settings for the program, recruitment and random assignment, and the treatment and control conditions.

3.2.1 Selection of Replication Grantees

The study design called for evaluating at least three replications of the model. In each of the replication sites, the services provided to youth in the intervention group had to be substantially different from the services provided to youth in the control group. In addition, grantees needed to be able to recruit enough youth over two years to meet the needs of the study. At the time of site selection, *¡Cuídate!* was being replicated by at least four TPP grantees. Since our evaluation design called for selection of at least three

⁶ A more detailed impact study design report can be found at <https://aspe.hhs.gov/basic-report/tpp-replication-study>.

⁷ The final impact report will answer a similar set of questions about program effects on intermediate outcomes and risk behavior after 18 months. It will also examine program impacts on pregnancy.

replications of each model, the choice of replication sites was constrained. Complicating the selection was the fact that most replications were not planned with the requirements of a rigorous evaluation in mind.⁸ It was apparent that one of the sites would not be able to build the sample of youth needed for the study over a period of two years. Combined with other considerations that could impede a strong test of the model, this led us to eliminate one of the four potential candidates. The three grantees selected were:

Community Action Partnership of San Luis Obispo County. Community Action Partnership, a non-profit agency founded in 1965 and based in San Luis Obispo, CA, provides a wide variety of programs and services to residents of San Luis Obispo County and 10 other California counties. Since 1977, the agency has provided comprehensive sexual health education programming in schools for youth ages 10–18. The agency also has its own reproductive health clinics, including teen-designed and peer-provided teen clinics.

La Alianza Hispana. Founded in 1970 and based in Boston, MA, La Alianza is a non-profit advocacy and service organization whose core programs address family mental health, public health and workforce education. The agency has worked with the Boston Housing Authority to provide information about HIV/STI and pregnancy prevention to young Latina women and, with other members of the Adolescent Trials Network (a collaborative of community-based organizations and healthcare providers, based at Boston Children’s Hospital), to reduce HIV infection rates among adolescents.

Touchstone Behavioral Health. This 30-year-old non-profit organization is based in Phoenix, AZ. Touchstone provides behavioral and mental health prevention and treatment programs and services to youth across Greater Phoenix. The agency has focused its prevention work on the Maryvale community, which has a predominantly Hispanic population. Before receiving the TPP grant, Touchstone had implemented a substance abuse prevention program and some limited sexual health programming in schools in this community.

Of the three replications, only La Alianza had some limited experience with *¡Cuídate!*. In 2009, the agency implemented a few program cycles in a small number of Boston public schools.

3.2.2 Site-Specific Program Designs

The three replications of *¡Cuídate!* shared important aspects of the program model. All three replications were required to implement the program with fidelity to the core elements of the model, and fidelity was assessed, monitored and reported to OAH at regular intervals by program staff. However, beyond these core elements, there were small variations in program design across the three sites. The Community Action Partnership replication added two sessions, to comply with the California requirements governing sex education—one on contraception and a second on STIs other than HIV. All three replications added a brief review of reproductive anatomy (to ensure that participants were aware of correct names for body parts). Touchstone also added a session on pregnancy prevention. Although the model recommends that

⁸ The 2010 TPP grant program included multiple funding ranges. All funded projects were expected to monitor and report on program implementation and outcomes through performance measures. Projects in the higher funding ranges (greater than \$1 million per year) were expected to be implemented in multiple sites within a targeted geographic area and were required to have an independent local evaluation. The *¡Cuídate!* replications selected for the study were in the lower funding range (less than \$1 million per year) and were not expected to have a rigorous local evaluation.

¡Cuidate! be delivered in groups of 6-10 youth, Touchstone received permission to deliver the program in larger groups of 20 to 25 students, with two health educators rather than one.

3.2.3 Settings for the Program

Grantees were selected for participation in the study between six and seven months after they had received the grant award. This meant that, for the most part, the grantees had recruited implementation settings prior to being chosen for the study. In two of the replications, *¡Cuidate!* was delivered in public school classrooms, as part of the regular school day. In San Luis Obispo County, the Community Action Partnership implemented the program in 10th-grade classrooms in three public high schools across the county. In Phoenix, Touchstone implemented the program in 8th-grade classrooms in eleven K–8 schools. In Greater Boston, La Alianza selected settings for the program that were more diverse: three public high schools (one traditional, one vocational-technical, and one charter school) and two community-based organizations (as part of a summer youth employment program, and a summer sports program). In two of the three schools the program was implemented in classrooms during the regular school day (9th- and 12th-grade classrooms in one school and mixed-age classrooms in the other).

3.2.4 Recruitment and Random Assignment

In each of the replication sites, individual students were recruited for the study. School staff identified classes or time slots in which *¡Cuidate!* would be offered (small groups of students would be pulled out of regular classes for the program).

Agency staff members who had been carefully trained by the Abt evaluation team conducted presentations to the identified groups or classes. In La Alianza, where the program was implemented in summer youth employment and sports programs, youth were recruited directly from the participant list. These presentations included information about the study procedures, a practical illustration of random assignment, and a description of the treatment and control conditions. The presentations were intended to personalize the study and help with recruiting.

The same agency staff distributed parent consent forms (both refusal and agreements) and study brochures and provided teachers with small incentives (\$5 gift cards) to offer students for the return of parent consent forms. They worked with individual teachers to gather the consent forms, and notified Abt study staff of students who had parent permission. Among those who returned consent forms, 89 percent had permission to participate in the study.⁹

Those with permission were invited to complete the baseline survey and included in the random assignment conducted by Abt. Participants were informed of their study assignment only after completing the baseline survey. Nearly all youth with parental consent completed the baseline survey. As Exhibit 3.1 shows, across the three sites combined, 2,198 students obtained parent permission and were randomized. Of those, 2,169 (98.7 percent) completed a baseline survey.

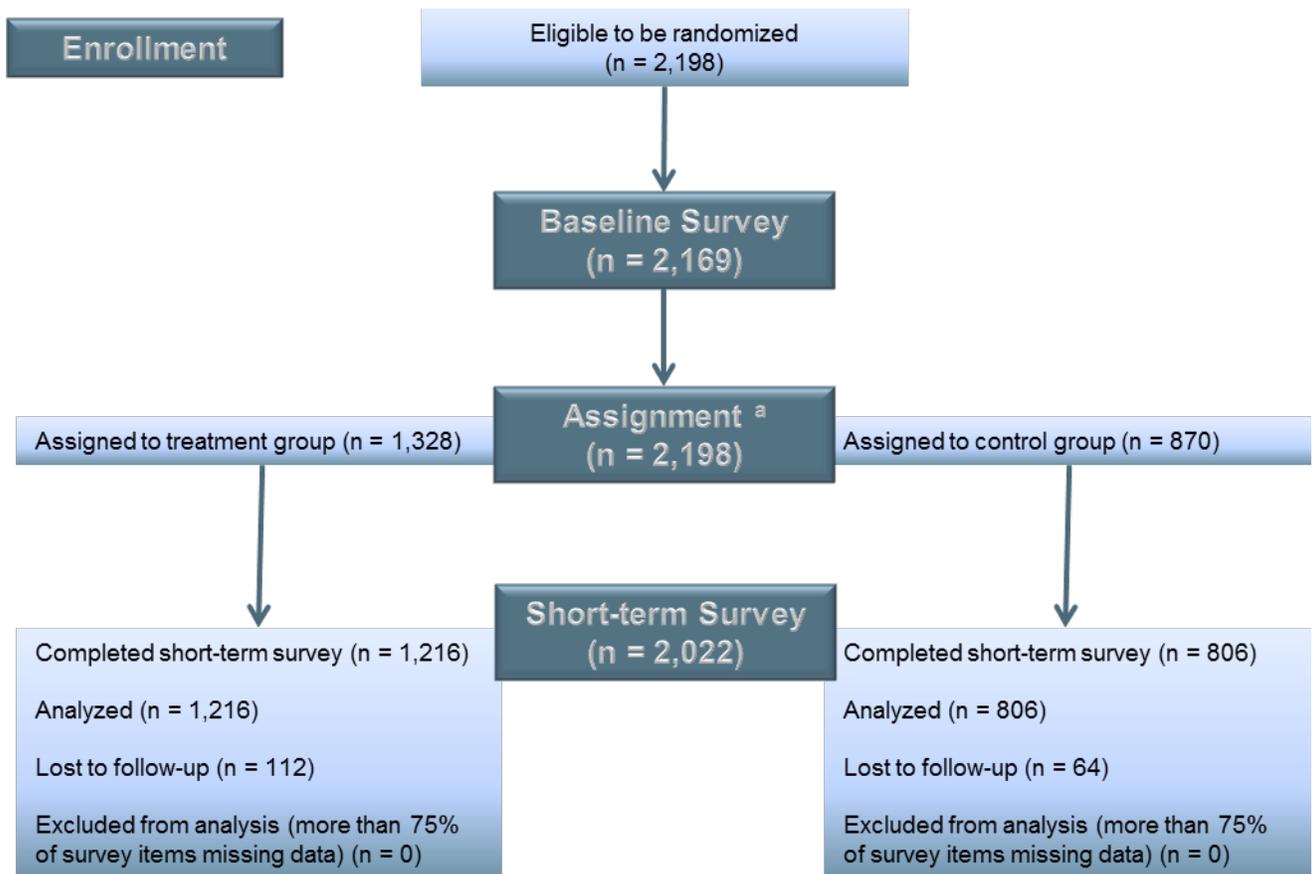
In each of the sites, students were randomly assigned within gender and class period blocks within school. In the two sites that offered *¡Cuidate!* in small groups during school or in other settings (La Alianza and

⁹ Data were not collected on students who declined to participate in the study. Therefore, it is not possible to assess similarities and differences between students who consented and those who did not.

Community Action Partnership), individual students were randomly assigned using a 2:1 ratio. That is, for every two students assigned to the treatment group, one student was assigned to the control group. In Touchstone, where the program was delivered in single-gender groups, random assignment was done by gender, using a 1:1 ratio.

Exhibit 3.1 shows how we arrived at the study’s analytic sample via the random assignment and survey completion processes, beginning with 2,198 eligible youth (i.e. those who had parent permission to participate).

Exhibit 3.1: Study Sample



^a A total of 29 participants did not take the baseline survey. Among them, 17 were assigned to the treatment group and 12 were assigned to the control group.

3.2.5 Treatment and Control Conditions

Across the three replications, members of the treatment group were offered all six sessions of *¡Cuidate!*, delivered by health educators who were members of the grantee staff, who were trained by the program developer and supervised and monitored continuously by grantee supervisory staff and external evaluators. Fidelity, quality and performance measures required by OAH were completed and reported by health educators, supervisory staff and external evaluators.

Beyond those important commonalities, there were differences in the treatment and control conditions across and within replication sites, as shown in Exhibit 3.2. The treatment condition varied across all three replication sites in: total number of sessions delivered; schedule for delivery; and age, gender, and size of groups. As for the control condition, in two of the three replication sites, students participated in a standardized activity (i.e., regularly scheduled PE classes in Community Action Partnership, and a healthy lifestyle curriculum in Touchstone). In the third replication site (La Alianza), where the program was delivered in a variety of settings, control group activities varied greatly.

Exhibit 3.2: Treatment and Control Conditions in the Three Replication Sites

Settings	Treatment Condition	Control Condition
Community Action Partnership of San Luis Obispo		
Tenth grade PE classes in three public high schools in San Luis Obispo County, CA	Number of sessions: 6 one-hour <i>¡Cuidate!</i> sessions plus 2 additional sessions, one on contraception, the other on STIs other than HIV Delivery schedule: Weekly Gender and size of groups: Small mixed-gender groups meet with a health educator	Regular weekly PE classes.
La Alianza Hispana		
<i>Multiple settings in Greater Boston, MA:</i>		
Ninth and 12th grade PE classes in a technical high school	Number of sessions: 8 forty-five-minute <i>¡Cuidate!</i> sessions Delivery schedule: Varied, daily or weekly depending on the grade Gender and size of groups: Small mixed-gender groups met with a health educator	Regular PE classes.
Mixed-age (9th–12th) health classes in a public charter school	Number of sessions: 6 one-hour <i>¡Cuidate!</i> sessions Delivery schedule: Weekly Gender and size of groups: Small mixed-gender groups met with a health educator	Regular weekly health classes.
After-school program in a public high school	Number of sessions: 6 one-hour <i>¡Cuidate!</i> sessions Delivery schedule: Daily Gender and size of groups: Small mixed-gender groups met with a health educator	No alternative provided for control group youth, but they were free to enroll in other after-school activities.
Early evening sports program serving a mix of ages in a community agency	Number of sessions: 3 two-hour <i>¡Cuidate!</i> sessions Delivery schedule: Daily Gender and size of groups: Small mixed-age mixed-gender groups met with a health educator	Control group youth were free to participate in sports activities offered.

Settings	Treatment Condition	Control Condition
Summer youth employment program in community agency, during a period of the day when youth are free to choose from a variety of enrichment activities	Number of sessions: 2 three-hour <i>¡Cuidate!</i> sessions Delivery schedule: Weekly Gender and size of groups: Small mixed-age mixed-gender groups met with a health educator	Control group youth were free to choose an activity (youth basketball was a frequent choice).
Touchstone Behavioral Health		
Eighth grade classes in eleven public K-8 schools in a single school district in Phoenix, AZ	Number of sessions: 6 one-hour <i>¡Cuidate!</i> sessions plus 2 additional sessions, one on pregnancy prevention, one to provide extra time Delivery schedule: Weekly Gender and size of groups: Single-gender groups of up to 25 students met with 2 co-facilitators	8 one-hour sessions of a healthy lifestyle curriculum. Single-gender groups met with two co-facilitators three times a week over a three-week period. Topics covered: self-esteem, body image, nutrition and exercise, stress management, decision-making skills, consequences of underage drinking, dangers of marijuana and other drugs, peer pressure.

3.3 Data Collection Strategies and Measures

3.3.1 Data Collection Strategy

To assess the impacts of the *¡Cuidate!* intervention, youth in the three replication sites were surveyed three times: at baseline, before the intervention began; six months after the baseline survey (short-term follow-up); and 18 months after the baseline survey (longer term follow-up). This report assesses outcomes using data from the 6 month survey; findings from the 18 month survey will follow in a later report. Each time, a web-based Audio Computer-Assisted Self-Interview (ACASI) system was used to capture and store survey responses and youth could choose to take the survey in Spanish or in English. At baseline, hard copies of the survey (in Spanish and English) were available as backup in case of computer or Internet failure.

In all three replication sites, study participants completed the baseline survey in group sessions in schools or other settings, using school computers where possible, or tablets dedicated to the study where it was not. Study staff oversaw the survey, and distributed survey incentives (gift cards) to students upon completion. Those who were absent on the day of the survey were contacted and a follow-up session was arranged.

For the first follow-up that is the subject of this report (6 months after baseline), only the web-based ACASI system was used. Text reminders were sent to all study participants (regardless of whether they had completed the baseline survey) before the survey went live and throughout the survey period.¹⁰ In some cases, before the survey window closed, field staff contacted participants and encouraged them to complete the survey independently on-line or helped them to access the survey. Youth could access the survey and complete it using personal tablets or computers, school or library computers, or even their smart phones. Gift cards were mailed to participants after completion of the survey.

¹⁰ Participants could complete the survey within a 3 month window.

As Exhibit 3.3 shows, 92 percent of eligible students completed the short-term follow-up survey (6 months after the baseline survey). There was almost no difference in the response rates of students in the treatment group compared with students in the control group.

Exhibit 3.3: ¡Cúidate! Short-term Survey Response Rate

	Eligible Participants			First Follow-up Completes Total		First Follow-up Completes Treatment		First Follow-up Completes Control	
	Total N	Treatment	Control	N	Percent	N	Percent	N	Percent
All Sites	2198	1328	870	2022	92.0%	1216	91.6%	806	92.6%
Community Action Partnership	530	313	217	490	92.5	287	91.7%	203	93.6%
La Alianza	728	463	265	692	95.1	439	94.8%	253	95.5%
Touchstone	940	552	388	840	89.4	490	88.8%	350	90.2%

3.3.2 Measures

The first follow-up survey collected information from students on a variety of topics, including questions that allow us to measure three sets of outcomes: 1) exposure to information about topics related to sexual risk behavior; 2) intermediate outcomes (i.e., factors that are believed to lead to behavioral outcomes); and 3) sexual activity and risk behavior. We briefly describe these measures here. A more complete description of these measures and the individual survey items that comprise them can be found in Appendix C. Here, Exhibit 3.4 summarized the outcome measures and their construction.

Exposure to Sexual Health Information. In the first follow-up survey, we asked students about their exposure to information about reproductive health and related topics. Students were asked if they had received information about any of a set of topics in the 12 months preceding the survey. Because the topics were distinct, we examined responses to individual survey questions rather than creating and analyzing a composite measure.

Intermediate Outcomes. Drawing on knowledge of the program’s theory of change and exploratory factor analysis, we constructed composite measures to assess four factors that potentially lead to behavioral outcomes: 1) **knowledge** of pregnancy risk and STI risk; 2) **attitudes towards protection and attitudes toward risky sexual behavior**; 3) **motivation to delay childbearing**; and 4) **condom negotiation and refusal skills**. For a fifth measure, we analyzed four single-item measures: 5) **intentions** to become sexually active in the immediate future and to use protection when sexually active.

Knowledge. We constructed two composite measures: *knowledge of pregnancy risk*; and *knowledge of STI risk*. The four items that make up the first measure and the 12 items that make up the second are all factual questions, testing students’ knowledge of the circumstances under which a woman can become pregnant (e.g., “A woman is protected from pregnancy the day she begins taking the pill”) and the effectiveness of condoms and other methods of birth control in preventing pregnancy, (e.g., “If birth control pills are used correctly and consistently, how much can they decrease the risk of pregnancy?”) as well as knowledge about STIs and their transmission (e.g., “You can’t get infected with HIV if you sex only once or twice without using a condom”). All items were scored 1 for a correct answer and 0 for any incorrect answer; scores were then averaged across the items that make up a measure and multiplied by 100 to indicate the percent of items answered correctly.

Attitudes. We constructed two composite measures of attitudes: *attitudes toward protection* and *attitudes toward risky sexual behavior*. For the 12 items that make up the first measure (attitudes toward protection), students were asked if they agreed or disagreed with statements such as “birth control is important to make sex safer.” Four response categories ranging from “strongly disagree” to “strongly agree” were scored from 1-4 and then scores for individual items were averaged, with higher values representing greater support for the use of protection. For the seven items that comprise the second measure (attitudes toward risky sexual behavior), students were asked if they agreed or disagreed with statements such as “It’s OK to have sex with someone on the first night you meet them.” Responses were scored 0 (disagree) or 1 (agree) averaged across the items and multiplied by 100 to indicate the percent of items agreed with, with higher scores representing higher levels of support for risky behavior.

Motivation. We constructed one composite measure, *motivation to delay childbearing*, which was made up of three survey items that asked students if they agreed or disagreed with statements such as “It is important for you to finish school before you have a child”. Four response categories ranging from “strongly disagree” to “strongly agree” were scored from 1 to 4 and then scores for individual items were averaged, with higher values representing greater motivation to delay childbearing.

Skills. We constructed two composite measures of skills: *condom negotiation skills* and *refusal skills*. The measure of *condom negotiation skills* comprised 7 items asking about respondents’ perceptions of their ability to obtain and negotiate the use of condoms with a partner (e.g., “If you were going to have sex, could you insist on using a condom even if your partner didn’t want to use one?”). Possible responses ranged from “I’m sure I could not” to “I’m sure I could” with the responses coded 1 to 4 and then averaged across items, with higher average scores representing greater certainty of condom negotiation skills.

The measure of *refusal skills* comprised six items probing respondents’ perceptions of their ability to refuse to engage in sexually risky behavior (e.g., “how sure are you that you would be able to say no to having sexual intercourse if neither you nor your partner had any form of birth control?”). Possible responses ranged from “I’m sure I could not” to “I’m sure I could” with responses coded 1 to 4. Scores for individual items were averaged, with higher average scores representing greater certainty of refusal skills.

Intentions. We included four single-item measures of students’ intention to engage in sexual activity in the year following the survey. The first item asked whether the respondent intended to have oral sex in the next year; the second asked if they intended to have sexual intercourse during the same period. The third asked if the respondent intended to use a condom if they were to engage in sexual intercourse in the next year; and the fourth item asked if the respondent intended to use birth control if they were to have sexual intercourse in the next year. Responses to each of the four items were scored 0 to 1 (1 for those responding that their intentions were “probably” or “definitely”).

Short-term Behavioral Outcomes. To address the study’s most important questions about the impact of the intervention, we identified eight measures in the domain of youth sexual behavior at the short-term follow-up: (1) *currently sexually active (in the last 90 days)*, (2) *sexual intercourse in the last 90 days*, (3) *oral sex in the last 90 days*, (4) *anal sex in the last 90 days*, (5) *initiation of sexual activity*, (6) *sexual intercourse without birth control in the last 90 days*, (7) *sexual intercourse without a condom in the last 90 days*, (8) *oral sex without a condom in the prior 90 days*, and (9) *anal sex without a condom in the last*

90 days. Sexual activity is defined differently across replication sites. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and/or anal sex. Youth were not asked about anal sex in Touchstone; therefore sexual activity refers to sexual intercourse or oral sex. All nine measures are single items, with yes/no answers.

Exhibit 3.4 summarizes the outcome measures and their construction.

Exhibit 3.4: Outcome Measures

Measure	Construction
INTERMEDIATE OUTCOMES	
Domain: Knowledge	
Knowledge of pregnancy risk	Continuous index: average of responses to four questions about circumstances in which it is possible to become pregnant and the extent to which contraceptive methods protect against pregnancy. Average scores multiplied by 100 range from 0-100 and represent the percent of the four questions answered correctly, with higher values representing more accurate knowledge.
Knowledge of STI risk	Continuous index: average of responses to 12 questions about STI transmission and prevention multiplied by 100. Scores range from 0–100 and represent the percent of the 12 questions answered correctly, with higher values representing more accurate knowledge.
Domain: Attitudes	
Attitudes toward protection	Continuous index: average of responses to 12 questions about attitudes towards using condoms and/or birth control during sex. Average scores range from 1 to 4 with higher values representing more positive attitudes toward using protection.
Attitudes toward risky sexual behavior	Continuous index: average score for seven binary items about the acceptability of risky sexual behavior. Average scores range from 0-1 with higher values representing more support for risky behavior.
Domain: Motivation	
Motivation to delay childbearing	Continuous index: average of three items about motivation to delay childbearing. Average scores range from 1 to 4 with higher values representing greater levels of motivation.
Domain: Intentions (in next 12 months)	
Intention to have oral sex in the next year	Single item scored 0 to 1, with 1 representing stronger intention.
Intention to have sexual intercourse in the next year	Single item scored 0 to 1, with 1 representing stronger intention.
Intention to use a condom if having sexual intercourse in the next year	Single item scored 0 to 1, with 1 representing stronger intention.
Intention to use birth control if having sexual intercourse in the next year	Single item scored 0 to 1, with 1 representing stronger intention.
Domain: Skills	
Refusal skills	Continuous index: average of responses to six questions about perceived ability to refuse to engage in risky sexual behavior. Average scores range from 1 to 4 with higher values representing greater certainty about refusal skills.
Condom negotiation skills	Continuous index: average of responses to 7 questions about perceived ability to obtain and negotiate the use of condoms. Average scores range from 1 to 4 with higher values representing greater certainty about condom skills.

Measure	Construction
BEHAVIORAL OUTCOMES	
Domain: Youth Sexual Behavior and Sexual Risk at the Short-Term Follow-up	
Currently sexually active (in last 90 days) ^a Sexual intercourse in last 90 days Oral sex in last 90 days Anal sex in last 90 days	Single items, yes (1) /no (0) answer.
Initiation of sexual activity	For those who were not sexually active at baseline, indicates whether they became sexually active between baseline and follow-up. Single item, yes (1) /no (0) answer.
Sexual intercourse without birth control in last 90 days ^a Sexual intercourse without a condom in last 90 days Oral sex without a condom in last 90 days Anal sex without a condom in last 90 days	Single items, yes (1) /no (0) answer.

^a Designated as a confirmatory outcome, as discussed in Section 3.4.5.

3.4 Analytic Approach

The impact analysis examines the extent to which *¡Cuídate!* affected each of the study’s outcomes. In testing for these effects, we use two-tailed hypothesis test procedures, because we do not want to rule out the possibility that the intervention might adversely affect one or more of the outcomes.

Our basic strategy for estimating program impacts is to compare the outcomes of treatment and control group members using a regression framework, in which we include baseline covariates to increase statistical precision (i.e., reduce the standard errors) of the impact estimates for a given sample size (Orr, 1999) and reduce attrition bias from missing data (see Puma et al., 2009).

3.4.1 Estimation of Impacts for the Full Sample

In this document, we report impact estimates that are pooled across the three *¡Cuídate!* replication sites. OAH’s requirements to define, measure, and adhere to fidelity to the program model means that each of the three replication sites implemented the same core program elements. The random assignment and data collection procedures were also the same across all sites. These design elements ensure that impact estimates pooled at the program level represent rigorous tests of a well-defined and consistently implemented program model.

For this evaluation of *¡Cuídate!*, individual sample members were randomly assigned to treatment or control conditions within randomization blocks, based on site, school, class period, and gender. For each outcome, we estimate a model that reflects this design, and has the basic structure of Equation 1.¹¹

$$(1) \quad Y_i = \beta_0 + \beta_1 T_i + \sum_{k=2}^{K+1} \beta_k X_{ki} + \sum_{m=1}^M \gamma_{0m} D_m + \varepsilon_i$$

In this model:¹²

¹¹ Because random assignment occurred at the individual level (not the classroom or school level) within randomization blocks, we estimated a one-level fixed-effects model that included a series of indicator variables representing each of the randomization blocks defined by site, school class period, and gender (Bloom, 2006, p.13).

Y_i is the outcome of interest (e.g. sexual intercourse without birth control) for the i^{th} individual in the m^{th} randomization block.

T_i is a dummy variable equal to 1 if individual i was assigned to the treatment group and 0 otherwise.

X_{ki} is the k^{th} baseline characteristic or covariate for individual i . These include baseline age, grade, race/ethnicity (black, white, Hispanic, other), living with biological parent/s, risk behaviors (smoking, alcohol use, marijuana use), baseline sexual activity (ever sexually active), baseline intentions to have oral sex and sexual intercourse, baseline pregnancy and STI risk knowledge, and baseline measure of the outcome when available.

D_m is the dummy variable representing the m^{th} randomization block. These block indicators reflect the fact that there were different treatment probabilities across blocks. Because random assignment blocks were constructed based on site, gender, class period, and school, the dummy variable also accounts for these factors.

ε_i is the usual random error term.

In this model, β_1 represents the average pooled impact of the program on the outcome. The coefficients on the covariates, β_k , reflect the relationship between the outcome measure and each of the covariates while controlling for others. It is important to note that this model specification treats randomization blocks (and thus sites) and the treatment effects as fixed as opposed to random, which is consistent with how the replication sites are chosen and how the results of the study will be interpreted.¹³

Equation 1 estimates the impact of access to *¡Cuidate!*. Because of the random assignment design, the crucial difference between the treatment and control groups is *access* to *¡Cuidate!* services: individuals in the treatment group had access to program services and possibly other sexual health education services available in the community (e.g., clinics), while control group members had access to only those other services in the community. In the evaluation literature, the estimate of the average impact of access is referred to as the intent-to-treat (ITT) impact parameter. It measures the average impact on treatment group members who had the **opportunity** to participate in the intervention, not the average impact on program group members who actually participated in the intervention. Very few participants did not receive at least one session of the intervention and, in fact, the majority of treatment group members received at least 75 percent of the program.¹⁴

¹² The analyses presented in this report used linear probability models for binary outcomes. A set of robustness analyses were conducted using multilevel logistic regression models and using multi-level linear models with heteroskedasticity robust standard errors for binary outcomes (Constantine et al, 2009, Gleason et al., 2010). There were no substantive differences in the inferences that results from any of the three modeling approaches.

¹³ Because replication sites were selected as a purposive sample, not randomly selected from a larger population of sites, we do not consider a random treatment effects model to be appropriate for drawing inferences in this sample (Schochet, 2008a, p. 70).

¹⁴ Class rosters were checked after random assignment results were communicated to the sites to verify the group status and determine if there were any cross-overs.

Finally, we report impact findings in tables showing the regression-adjusted treatment group mean (calculated as the sum of the control group mean and the regression adjusted treatment effect), the unadjusted control group mean, and the regression-adjusted treatment effect (and p-value). For binary outcomes (e.g., condom use), we report impacts as percentage point differences between the treatment and control group means. For all other outcomes, we show impact estimates in their original metric and additionally convert impact estimates to effect sizes (by dividing the impact estimate by the control group standard deviation) and report these in a separate column.

3.4.2 Site-level Analyses

In addition to estimating impacts pooled across the three replication sites, we estimate impacts for each site separately and test for differences in impact across the three sites. We implement these analyses by including treatment by site interaction terms in the model (i.e., Equation 1) and testing for the joint significance of the interaction terms. When statistically significant differences in impact are found across sites for one or more outcomes, we discuss these differences.¹⁵ Site-specific impact estimates for all outcomes are presented in Appendix A.

3.4.3 Subgroup Analyses

In addition to the overall pooled impacts and site-level impacts, we estimate impacts for key subgroups of participants (based on age, gender, race/ethnicity, and sexual experience at baseline) and test for differences between subgroups, to better understand what works for whom. We implement subgroup analyses by including subgroup indicators and treatment by subgroup interaction terms in the model (i.e., Equation 1) and testing for significance of the interaction term.¹⁶

To guard against potential over-interpretation of results among the very large number of subgroup estimates, we only present impact estimates for individual subgroups when there is a statistically significant difference between subgroups; for example, the impact would be presented for the subgroup of boys only if there were a statistically significant difference in impacts between boys and girls (see Appendix B).

3.4.4 Handling Missing Data

We used monetary incentives and intensive tracking to achieve the maximum possible response rate for both treatment and control groups, and have so far achieved very high response rates in each of the replication sites (see Exhibit 3.3 above).

¹⁵ The purpose of testing for differences across sites before discussing results in the main text is to guard against over-interpretation of spurious findings, some of which would be expected by chance in such a large group of outcomes. The basic idea behind the strategy of discussing site-specific impacts only when differences are found is that it is only credible to report an impact in one site—but not in another—if there is a significant difference between the sites. The site-specific results in Appendix A are not adjusted for multiple comparisons and any significant findings reported there should be interpreted with caution.

¹⁶ For the treatment-by-race/ethnicity interaction, a 3 degrees-of-freedom F test was used.

We use case deletion for the few instances of missing outcome data (Puma et al., 2009). Dummy-variable adjustment is used in regression models to account for missing covariates. In the dummy variable adjustment method, missing covariate values are set to a constant and indicators (or dummy variables) for such values are added to the impact analysis model (Puma et al., 2009).

3.4.5 Addressing Multiple Comparisons

Ongoing developments in the statistical analysis of the results of randomized trials emphasize that conventional statistical tests and confidence intervals apply to a single outcome. When analysts look over multiple outcomes for any statistically significant finding, the appropriate critical t-values are much higher; i.e., effects that appear to be statistically different from zero are not truly different from zero. In the literature, this is known as the problem of “multiple comparisons.” Current guidance on how to approach this multiple comparison problem recommends distinguishing two categories of analyses (Schochet, 2008b). One—called “confirmatory tests”—includes a small number of critical outcome domains for which it is important to adjust error probabilities for multiplicity. Confirmatory analysis uses a high standard of evidence for deciding if an intervention has had its intended effect, in order for its findings to be considered conclusive rather than merely suggestive. A second category includes “exploratory tests” for which there is generally higher tolerance of errors and for which multiplicity adjustments may or may not be made.

For this report, the impact analysis team pre-specified a multiple comparisons strategy that spans the two reports and includes confirmatory and exploratory analyses. The confirmatory analysis seeks convincing evidence that *¡Cuidate!* improved participants’ behavioral outcomes past the end of the program. Before analyzing data, the team pre-specified a small number of outcomes in three “domains,” or sets of similar constructs, as part of the confirmatory analysis. The three confirmatory outcome domains are: *youth sexual behavior at the short-term follow-up; recent sexual behavior at the longer-term follow-up; and pregnancy.*

To control for multiple comparisons within each of the domains, we apply a formal multiple comparisons correction (in particular, a Benjamini-Hochberg correction, as described in Appendix G of the What Works Clearinghouse Procedures and Standards Handbook, version 3.0, which controls for the false positive rate by adjusting P-value thresholds). The correction does not affect the p-values that appear in tables of results, but it does change the interpretation of statistical significance. In particular, it raises the bar for rejecting the null hypothesis.

Two outcomes in this interim report, **currently sexually active** (engaged in sexual intercourse, oral sex, or anal sex in the last 90 days) and **sexual intercourse without birth control** (engaged in sexual intercourse without a condom or other birth control in the last 90 days), were pre-specified as key outcomes in one of the study’s three confirmatory outcome domains: youth sexual behavior and sexual risk at the short-term follow-up. These findings stand alone as rigorous evidence of *¡Cuidate!*’s short-term behavioral impact. The other two confirmatory outcome domains, “Recent sexual behavior at the long-term follow up” and “Pregnancy” will be analyzed in the final report, along with the findings presented here.

The exploratory analysis encompasses all other outcomes and research interests in the interim report, e.g. impacts on intermediate outcomes and other behavioral outcomes. Given the large number of hypothesis tests that constitute the exploratory analysis, some false positive findings are to be expected. We do not make formal adjustments for multiple comparisons when reporting on statistical significance. However, to

aid in interpretation, we specify the number of tests that were conducted (within and across domains) and the number of false rejections that would be expected given the number of tests if there were no impact of treatment.

4. Results

The study is designed to determine whether *¡Cuidate!* helps young people develop the knowledge, attitudes and skills to act in ways that ultimately protect them from the potential consequences of sexual risk behavior (i.e. STIs, HIV and early pregnancy). The program, when delivered with fidelity, is intended to provide information and affect potential intermediate outcomes such as: knowledge and understanding of reproductive health and avoidance of sexual risk; attitudes toward using protection; motivation to avoid pregnancy; intentions to become sexually active and use protection; and skills needed to avoid sexual risk. The ultimate goals are reduced rates of unprotected sexual activity and unplanned pregnancy.

The short-term findings discussed here suggest that:

- *¡Cuidate!* was implemented as intended.
- It was effective in increasing knowledge, attitudes and perceived negotiation skills.
- However, after six months, these changes did not lead to improvements in reported sexual risk behaviors.

In this section, we expand on our conclusion that the program was indeed implemented with fidelity across replication sites, discuss findings for the full study sample and for individual sites, as well as any important findings for specific subgroups of youth (age, gender, race/ethnicity, and sexual experience at baseline).

In addition to the exhibits in this section, tables documenting the site level analyses can be found in Appendix A and the corresponding tables documenting subgroup analyses can be found in Appendix B.

4.1 Program Implementation

As we noted in Section 1.3, a separate report (in preparation) will provide a detailed account of the implementation of *¡Cuidate!* in the three replication sites. The implementation report will serve two important purposes: (1) to help explain the findings of the Impact Study; and (2) to offer lessons learned to help those planning to use *¡Cuidate!* in the future.

What we have learned from the implementation study that is directly relevant for this report is that the program was generally well implemented across the three replications. The three grantees hired staff with appropriate background experience and skills to deliver the program; all received training approved by the developer; the program was implemented with fidelity to the core elements and without modifications that threatened those core elements; and attendance was generally strong.

4.1.1 Staff Hiring and Training

Although the grantees gave the staff who administered the program different titles (e.g., prevention specialist vs. health educator), they were consistent in the types of experience and skills that they looked for when hiring. Experience working with adolescents and bilingualism, or a Hispanic background, were considered essential. In addition, and given equal emphasis, was comfort in addressing sexual health issues. (One of the grantees required candidates to give a presentation on adolescent sexual health to staff as part of the screening process.) Ultimately, staff hired to deliver the program brought a variety of skills and experience to their roles—their backgrounds included: clinical work in reproductive health; tutoring at-risk youth; adolescent mental health; dating violence and sexual assault prevention; and HIV/AIDS

outreach work. All of the staff received the official training provided by the curriculum distributor and approved by the developer. Grantees offered additional training and encouraged staff to attend training sessions offered by OAH, as well as by state or local agencies and institutions. Staff retention was high.

4.1.2 Implementing the Program with Fidelity

As part of the TPP program, OAH stipulated that grantees maintain fidelity to the core components of the program model, and provided guidance on making minor adaptations (all of which had to be approved by OAH before they could be implemented). There was an accompanying requirement that grantees develop a fidelity monitoring plan that would allow for monitoring of implementation and continued adherence to the core program model.

For *¡Cuidate!*, fidelity monitoring log templates were provided by the developer for assistance in collecting this information. Health educators were required to complete a fidelity log for each session delivered. In addition, OAH provided observation protocols, to be used by supervisory staff on a regular schedule that allowed an assessment of the **quality** of the sessions. Data from the logs and observations were aggregated and used by program supervisory staff to identify areas where improvement was needed. Aggregate data were delivered to OAH every six months and summarized to provide a basis for subsequent discussions between OAH program officers and the grantees. All of these activities were intended to guide implementation and ensure a degree of uniformity across grantees implementing the same program model.

Although there were variations in the program design for each replication site, as we described earlier, this variation did not affect implementation of the core elements of the program model. Each of the replication sites successfully delivered the intervention to students with fidelity to the program model. Nevertheless, grantees discovered that they needed to develop strategies to address implementation challenges. As we noted earlier, the *¡Cuidate!* curriculum used by all three grantees paid scant attention to pregnancy prevention. The three grantees dealt with this gap in different ways—two grantees added a unit, the third did not supplement the curriculum. In two of the three replications, health educators faced the challenge of making the curriculum relevant to non-Hispanic youth who chose to enroll in the study. In most settings, but especially in school settings, targeting a specific ethnic population is not permitted, for obvious reasons. And the developer, who included non-Hispanic youth in her test of the program, believes that the concepts are relevant to youth of any background. As a consequence, grantee staff delivering the curriculum had to find ways to translate the cultural references and concepts to make them relevant to non-Hispanic youth.

4.1.3 Participant Attendance and Engagement

Grantees were required to collect and report youth program attendance (by session) using attendance logs. In all three replications, a majority of students received at least three-quarters of the sessions offered. The numbers differed slightly by site, with Touchstone having the highest attendance and La Alianza having the lowest.¹⁷ In Touchstone almost everyone attended all of the eight sessions.

¹⁷ In La Alianza, attendance was lower where the program was implemented outside the regular school day or in non-school settings (one after-school setting and two community settings).

Abt's independent observations and focus group sessions with students suggest that they actively participated and acquired new information in the sessions.

4.2 Sample Characteristics

4.2.1 Study Sample

Baseline characteristics of the overall *¡Cuídate!* sample and for each replication site are presented in Exhibit 4.1. At baseline, youth in the study sample were, on average, 14-and-a-half years old. However there was considerable variation across the replication sites; in Touchstone, where the program was implemented only in 8th grade classrooms, the average age of students was about 13 years—more than one year less than the average for the combined sample.

Female teens constituted more than half of the sample. More than 70 percent were Hispanic, almost 20 percent were White and the remaining 10 percent were divided almost equally between Black and Other race (which includes Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, or Multiracial).

The site-specific profiles of youth participating in the study differ from each other in several ways. Given the program's focus on Latino culture, perhaps the most important is the difference in the ethnic mix of the samples: In Community Action Partnership, just over half were Hispanic and more than one-third were White, compared with significantly higher proportions of Hispanic youth and much lower proportions of White youth in the other two replication sites.

Across all three replication sites, more than 90 percent of youth lived with one or both biological parents. Almost half reported feeling very close to and cared for by their fathers; 60 percent reported they felt close to and cared for by their mothers.

Almost half of the sample had ever used alcohol and more than a quarter had used marijuana. Less than 20 percent had ever smoked cigarettes.

Not quite half of the students (48 percent) in the overall sample demonstrated an accurate understanding of pregnancy risk, while a somewhat smaller proportion (39 percent) understood STI risks. Across all three sites, the majority of students reported supportive attitudes toward using protection. Overall, less than one-quarter of the sample expressed an intention to engage in oral sex in the next 12 months; almost one-third (31 percent) intended to have sexual intercourse in the same period. The younger students in the Touchstone sample were significantly less knowledgeable about pregnancy and STI risks and significantly fewer of them expressed intentions to engage in sex in the next 12 months compared with students in the other two sites. Across the sample, students reported strong intentions to use protection if they were to have sexual intercourse in the next 12 months.

There were substantial and significant differences among the sites in the extent to which students had engaged in sexual activity and sexual risk behavior before they entered the study. While less than a quarter of the overall sample had ever been sexually active, and 17 percent were currently sexually active, the younger Touchstone students were strikingly less sexually experienced: just 7 percent had ever been sexually active and less than 4 percent were currently sexually active. Students in La Alianza consistently reported the highest levels of sexual activity. This same pattern repeated for sexual risk behaviors. The proportions of students who had engaged in unprotected sex were consistently lowest in the Touchstone

sample and highest in the La Alianza sample, where students were, on average, about a year older than the average age of the sample as a whole.

Just over half of the students in the sample reported receiving information on birth control methods, and fewer reported receiving information on how to obtain birth control in the year before the study began. Nearly three-quarters of participants reported receiving some information on STIs, probably from one or more health classes providing mandated lessons on HIV.

Youth in the Touchstone sample were much less likely than youth in the other sites to have been exposed to information that would help to protect them against sexual risks. Because youth in Touchstone were all in the 8th grade, usually the last grade in middle or intermediate school, they may not have had health classes and almost certainly had not had classes in which the topic of sexual health was addressed.

Exhibit 4.1: Baseline Characteristics of the Analytic Sample by Site

Measure	Community Action Partnership	La Alianza	Touchstone	<i>¡Cuidate!</i> Overall	p-value for the Test of Differences across Sites ^a	
Demographic characteristics						
Age (years)						
Mean	15.00	15.58	13.19	14.45	0.000	***
Grade	9.99	10.17	8.00	9.23	0.000	***
Gender						
Female	61.02	56.50	51.07	55.34	0.002	**
Race/ethnicity^b						
Hispanic	52.04	80.20	75.00	71.22	0.000	***
Black	1.02	6.07	5.83	4.75	0.000	***
White	35.92	11.13	12.62	17.75	0.000	***
Other	11.02	2.60	6.55	6.28	0.000	***
Family structure and relationships						
Lives with biological parent/s	93.00	89.96	94.76	92.65	0.002	**
Feels very close to and cared for by father	42.98	36.15	54.12	45.28	0.000	***
Feels very close to and cared for by mother	53.83	57.38	65.13	59.64	0.000	***
Risk behaviors						
Ever smoked cigarettes	26.90	20.09	12.16	18.53	0.000	***
Ever used alcohol	62.70	57.04	29.53	47.22	0.000	***
Ever used marijuana	38.60	31.72	13.15	25.84	0.000	***
Knowledge, attitudes and intentions						
Knowledge of pregnancy risk	60.55	54.24	36.07	48.29	0.000	***
Knowledge of STI transmission	52.25	44.50	26.47	38.96	0.000	***
Attitudes toward protection (1 = least supportive, 4 = most supportive)	3.11	3.12	3.02	3.07	0.000	***

RESULTS

Measure	Community Action Partnership	La Alianza	Touchstone	<i>¡Cuidate!</i> Overall	p-value for the Test of Differences across Sites ^a	
Intentions to have oral sex in the next 12 months	31.04	34.04	12.12	24.17	0.000	***
Intentions to have sexual intercourse in the next 12 months	36.46	44.89	15.32	30.53	0.000	***
Intentions to use a condom if they were to have sexual intercourse in the next 12 months	95.43	93.12	92.69	93.51	0.137	
Intentions to use birth control if they were to have sexual intercourse in the next 12 months	93.53	90.14	92.02	91.75	0.114	
Sexual Behavior^c						
Ever sexually active	31.61	41.31	7.18	24.76	0.000	***
Currently sexually active in the last 90 days	21.53	29.60	3.79	16.90	0.000	***
Sexual intercourse in the last 90 days	16.74	25.82	3.30	14.25	0.000	***
Oral sex in the last 90 days	16.15	21.34	2.20	12.12	0.000	***
Anal sex in the last 90 days	2.28	5.38		4.08	0.009	**
Sexual Risk^c						
Sexual intercourse without a condom in the last 90 days	8.88	15.07	1.10	7.76	0.000	***
Oral sex without a condom in the last 90 days	15.11	19.10	1.34	10.75	0.000	***
Anal sex without a condom in the last 90 days	1.86	3.14		2.60	0.180	
Sexual intercourse without birth control in the last 90 days	5.99	6.72	0.86	4.11	0.000	***
Baseline exposure to program information^d						
Relationships or marriage	87.42	74.23	75.40	77.93	0.000	***
Abstinence from sex	84.60	66.08	39.00	59.64	0.000	***
Birth control methods	79.01	64.56	28.45	53.29	0.000	***
Where to obtain birth control	70.72	58.74	20.42	45.93	0.000	***
Sexually transmitted diseases/infections	94.65	82.75	53.38	73.63	0.000	***
How to talk with partner about sex and birth control	66.26	56.98	22.15	45.01	0.000	***
How to say no to sex	79.01	70.31	56.20	66.63	0.000	***
How babies are made	94.86	86.20	77.75	84.84	0.000	***

Source: Baseline survey administered prior to randomization.

Notes: Results in this table are based on 1,940 - 2,022 respondents (for *¡Cuidate!* overall) who provided valid survey responses to relevant items except for the items measuring how close the respondent feels to their mother (n=1,930) and father (n=1,800), and anal sex (n=1,152). Values shown are percentages unless otherwise indicated. The items that compose measures of attitudes toward risky sexual behavior, motivation to delay childbearing, refusal skills, and condom negotiation skills were not asked at baseline.

^a Test results from an analysis of variance testing the null hypothesis that the means of the variable indicated in the row are equivalent among the three sites.

^b Racial ethnic categories are Hispanic, Black non-Hispanic, White non-Hispanic, and other race non-Hispanic, where other is defined as Asian, American Indian or Alaska native, native Hawaiian or other Pacific Islander, multiracial or undisclosed race.

^c Sexual activity is defined differently across grantees. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Touchstone.

^d Questions refer to information received in the 12 months prior to the survey administration.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

4.2.2 Comparability of the Treatment and Control Groups at Baseline

Although the characteristics of study participants differed significantly across the three replication sites (reflecting the differences in youth populations in those sites), there were few significant differences between those assigned to the treatment group and those assigned to the control group.

Baseline treatment-control differences were estimated using a series of models with the same structural components as the impact model in Equation 1 (i.e., the same randomization block indicators and treatment group indicator), but where, in each model, one of the baseline characteristics in Exhibit 4.1 served as the dependent variable, and where the other covariates used in the impact model were omitted. In this approach, the coefficient for the treatment indicator is the treatment-control difference on the pre-test measure. There were few significant differences between the two groups on any of the measures (See Appendix Table D.1); variables for which there were differences were subsequently included in the impact models as covariates.

4.3 Program Impacts on Exposure to Sexual Health Information

In each of the three replication sites, the *¡Cuidate!* curriculum represented a way to provide primarily Hispanic youth with sexual health information that may or may not be available from other sources. Each of the replications sought to supplement what youth typically get in schools (usually two sessions in a high school health class) and address a perceived lack of services and information.

We anticipated that, after participation in *¡Cuidate!*, youth in the treatment group would be more likely to report exposure to sexual-health-related information than their counterparts in the control group. As expected, *¡Cuidate!* had a statistically significant and positive effect on students' exposure to information about sexual health topics. Six months after the program began:

- There were large and significant effects on students' exposure to information about birth control methods, where to obtain birth control and how to talk to a partner about birth control. compared with students in the control group.
- Participants assigned to *¡Cuidate!* were more likely to report receiving information about STIs, abstinence, how babies are made, and how to say no to sex (Exhibit 4.2).¹⁸ Focus group interviews with program participants indicated that the condom information (and demonstration) was information that they had not received elsewhere.

¹⁸ There were significant, positive effects on eight out of eight measures exposure to program information (Exhibit 4.2). In the absence of a true program impact, with eight tests and a significance criterion of $p < 0.05$ the expected number of findings that would be significant by chance alone is less than one.

Site-Level Differences. In each of the sites, significantly more program participants reported being exposed to information on sexual-health-related topics than their control group counterparts.

- The magnitude of the differences was greatest in Touchstone (See Appendix Table A1). As mentioned earlier, two of the replications supplemented the curriculum with information specifically about birth control to address both gaps in the curriculum content and state health requirements. Without these additions, it is likely that the effect on exposure to information about birth control would not have been as strong.

Exhibit 4.2: Short-term Impacts on Exposure to Sexual Health Information

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value
Percentage of respondents that reported receiving information on the following topics:				
Relationships or marriage	89.33	81.89	7.44***	0.000
Abstinence from sex	83.12	67.00	16.12***	0.000
Birth control methods	80.64	59.53	21.11***	0.000
Where to obtain birth control	74.65	51.24	23.41***	0.000
Sexually transmitted infections	91.46	74.63	16.83***	0.000
How to talk with partner about sex and birth control	84.88	57.52	27.37***	0.000
How to say no to sex	89.71	74.35	15.36***	0.000
How babies are made	92.20	87.45	4.75***	0.000

Source: Follow-up survey administered six months after baseline.

Notes: Questions refer to information received in the 12 months prior to the survey administration. Appendix C provides detailed information on measures. Results in this table are based on 2,015 - 2,021 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

Subgroup Differences. There were significant differences in exposure to program messages by specific subgroups defined by: sexual experience at baseline; student age; and gender.

Sexual experience at baseline. Exposure to information varied by whether or not participants had ever been sexually active at baseline. There were significant program impacts on exposure to information about relationships and how babies are made among those who were and were not sexually experienced at baseline, but the impacts were stronger among those who were sexually experienced at baseline. There was a significant program impact on exposure to information about talking to a partner about sex or birth control for program participants overall, but the impact was greater for those who were not sexually experienced at baseline.

Age. Program effects on exposure to sexual-health-related topics differed by age, with a greater percentage of younger treatment group members reporting exposure to information than their control group counterparts. Program effects were smaller (but still significant) for older participants.

Gender. A greater percentage of males and females in the program group received information on how to say no to sex compared with those in the control group, but the difference was greater for males (see Appendix Table B.1).

4.4 Intermediate Outcomes

¡Cúidate!'s theory of change (see logic model in Exhibit 2.2) specifies intermediate outcomes that the model predicts will influence behavior—namely, **knowledge** and understanding of sexual risk behaviors and their prevention or avoidance; **attitudes** toward protection sexual risk behaviors; **motivations** and **intentions** to engage in sexual activity; communication, refusal and negotiation **skills**. We find evidence that:

- *¡Cúidate!* had positive impacts on knowledge, attitudes, and some skills.
- There were no program effects on motivation or on intention to engage in sexual behaviors in the following year.

4.4.1 Knowledge

Knowledge of Pregnancy Risk. The effect of *¡Cúidate!* on knowledge of pregnancy risk was assessed using a composite measure that combines four survey items on topics such as the effectiveness of condoms and birth control in preventing pregnancy. Compared with control group students, treatment group students scored significantly higher on the composite measure, and were significantly more likely to give the correct answers for each of the individual items.¹⁹ Exhibit 4.3 shows the findings for the composite measure and the individual items.

Knowledge of STI Risk. *¡Cúidate!* also had large and statistically significant positive impacts on a composite measure and on individual survey items measuring student knowledge of STI risk (prevention, transmission and treatment). Students in the treatment group were more likely to answer correctly questions about the effectiveness of birth control and condoms in preventing STIs. There were also large impacts on students' understanding of the transmission of STIs, and the consequences of sexual activity.

Of the items in the composite measure, the one on which the program had no impact was knowledge of the HPV vaccine. Fewer than half of the students in both groups were aware that a vaccine existed. In this context, it is worth noting that, although the program had a significant impact on students' understanding of the effectiveness of condoms in preventing both gonorrhea and HIV (and the ineffectiveness of birth control pills to prevent gonorrhea), even in the treatment group, the percentages who answered these questions correctly were quite low.²⁰ Exhibit 4.3 shows the findings for the composite measure and the individual survey items.

¹⁹ There were significant, positive effects on five out of five measures of knowledge of pregnancy risk (Exhibit 4.2). In the absence of a true program impact, with five tests and a significance criterion of $p < 0.05$ the expected number of findings that would be significant by chance alone is less than one.

²⁰ There were significant, positive effects on twelve out of thirteen measures of STD risk (Exhibit 4.2). In the absence of a true program impact, with thirteen tests and a significance criterion of $p < 0.05$ the expected number of findings that would be significant by chance alone is about one.

Site-Level Differences. Program impacts on knowledge about pregnancy and STI risks were largest for participants in Touchstone, but were positive and significant for both pregnancy and STI risk in Community Action Partnership, and for STI risk in La Alianza (See Appendix Table A.2).

Exhibit 4.3: Short-term Impacts on Knowledge of Pregnancy Risk and STI Risk

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value
Knowledge of pregnancy risk (percent of items respondents answered correctly) ^c	67.07	60.95	6.12***	0.000
Percent of respondents correctly answering each item				
Used correctly, how much can birth control pills reduce pregnancy risk?	60.98	54.71	6.26**	0.004
Used correctly, how much can condoms reduce pregnancy risk?	65.41	57.07	8.34***	0.000
A couple that has had unprotected sex and not gotten pregnant does not have to worry about getting pregnant.	84.56	79.78	4.78**	0.005
A woman is protected from pregnancy the day she begins taking the pill.	57.31	52.23	5.07*	0.020
Knowledge of STI risk (percent of items respondents answered correctly) ^d	63.67	53.01	10.66***	0.000
Percent of respondents correctly answering each item				
Once you are infected with HIV you are infected for life	81.85	69.23	12.62***	0.000
There is a vaccine to prevent girls from getting HPV	43.66	42.68	0.98	0.662
All STD/STIs can be cured by taking medicine	75.15	59.93	15.23***	0.000
A person with an STD/STI who looks and feels healthy cannot transmit the infection to others	82.55	75.31	7.23***	0.000
Some STDs/STIs put you at greater risk of HIV	71.63	61.17	10.46***	0.000
About 1 out of 4 sexually active teens gets an STD/STI every year	67.43	59.43	8.00***	0.000
You can get an STD/STI from having oral sex	73.63	62.03	11.60***	0.000
Used correctly, how much can condoms decrease the risk of HIV?	54.12	43.18	10.95***	0.000
You can get HIV even if you unprotected sex only 1 or 2 times	71.61	60.30	11.31***	0.000
Used correctly, how much can condoms decrease the risk of gonorrhea?	34.67	27.67	7.01***	0.001
Used correctly, how much can birth control pills decrease the risk of HIV?	58.58	39.33	19.25***	0.000
Used correctly, how much can birth control pills decrease the risk of gonorrhea	48.36	35.86	12.50***	0.000

Source: Follow-up survey administered 6 months after baseline.

Notes: Results in this table are based on 2,022 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c Score based on the four items below. Values represent the average percent of items answered correctly by respondent for each group. Alpha coefficient = 0.50.

^d Score based on the 12 items below. Values shown represent the average percent of items answered correctly by respondent for each group. Alpha coefficient = 0.69.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Subgroup Differences. The impacts of the program on the composite measure of knowledge of pregnancy risk were different by sexual experience, age and gender.

Sexual Experience. Both program participants who were never sexually active at baseline as well as those who were experienced, demonstrated significantly greater knowledge of STI risk than their counterparts in the control group, but the impacts were considerably larger for those who had never been sexually active at baseline.

Gender. The impacts of the program were different for males and females. Impacts were large and statistically significant for females while for males the effect approached statistical significance (See Appendix Table B.2).

Age. Program effects on knowledge differed by age as well. Younger program participants were significantly more likely to answer correctly questions about pregnancy and STI risk than their control group counterparts. While program effects were significant or approached statistical significance for older program participants, the impacts on younger participants were more than twice as large as those for older participants.

4.4.2 Attitudes

Attitudes Toward Protection. *¡Cuidate!* had modest but statistically significant impacts on students' attitudes toward using birth control or condoms (SES 0.24). For the composite measure on attitudes toward protection (comprising both birth control and condoms), students in the treatment group had more positive (and protective) attitudes. On a scale ranging from 1 to 4 with higher values indicating more positive attitudes, the mean for the treatment group was 3.24. There were positive and statistically significant differences on most of the items within the scale (Exhibit 4.4).²¹

Site-Level Differences. Program impacts differed by replication site, with significant positive impacts in Touchstone (See Appendix Table A.3).

Subgroup Differences. There was a significant difference on program impacts by age. Younger program participants had more positive attitudes toward using protection than younger participants in the control group. The differences for older program participants were not significant (See Appendix Table B.2).

²¹ There were significant, positive effects on ten out of 13 measures of attitudes toward protection (Exhibit 4.3). In the absence of a true program impact, with 13 tests and a significance criterion of $p < 0.05$ the expected number of findings that would be significant by chance alone is about one.

Exhibit 4.4: Short-term Impacts on Attitudes Toward Protection

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value	Effect Size ^c
Attitudes toward protection ^d	3.24	3.14	0.10***	0.000	0.24
Birth control pills should always be used if a person your age has sexual intercourse	3.27	3.29	-0.02	0.576	-0.03
Birth control is too much trouble to use (reverse)	3.21	3.09	0.12***	0.001	0.15
Birth control is pretty easy to get	2.80	2.74	0.05	0.162	0.06
Birth control is important to make sex safer	3.27	3.25	0.02	0.623	0.02
Birth control has too many side effects (reverse)	2.75	2.60	0.15***	0.000	0.21
Using birth control is morally wrong (reverse)	3.33	3.22	0.11***	0.000	0.16
Condoms are too much trouble to use (reverse)	3.35	3.23	0.12**	0.001	0.14
Condoms are pretty easy to get	3.27	3.19	0.07*	0.035	0.09
Condoms are important to make sex safer	3.66	3.52	0.14***	0.000	0.21
Using condoms means you don't trust your partner (reverse)	3.48	3.35	0.13***	0.000	0.17
Using condoms is morally wrong (reverse)	3.57	3.40	0.17***	0.000	0.25
Condoms decrease sexual pleasure (reverse)	2.87	2.71	0.16***	0.000	0.19

Source: Follow-up survey administered 6 months after baseline.

Notes: Results in this table are based on 1,946 - 2,022 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c The effect size is the standardized effect size of the difference, which is the "treatment effect" divided by the pooled standard deviation of the treatment and control groups.

^d This construct averages responses to 12 items (shown in table) on attitudes towards condoms and birth control. Possible values range from 1 to 4 with higher values indicating more positive attitudes toward protection (alpha coefficient = 0.79).

* p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

Attitudes Toward Risky Sexual Behavior. ¡Cuidate! had no statistically significant impacts on the composite measure or on any of six individual items measuring attitudes toward risky sexual behavior (Exhibit 4.5). Six months after baseline, students in both the treatment and control groups rejected the view that risky behavior was acceptable. On the composite measure that averaged responses to seven items, with values of 100 indicating the highest (i.e., agreement with all items) and 0 indicating the lowest levels of support for risky behavior, the mean was quite low and almost equivalent for both groups (3.1% and 3.3%). For each of the individual items, the means of the two groups were also very similar. This suggests that at minimum, students know the socially acceptable answer.

There were no significant differences on the composite measure of attitudes toward risky sexual behavior by replication site (Appendix Table A.4) or across subgroups (Appendix B).

Exhibit 4.5: Short-term Impacts on Attitudes Toward Risky Sexual Behavior

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value
Attitudes toward risky sexual behavior (ranked importance) ^c	3.12	3.33	-0.21	0.692
Percent of respondents agreeing with each item				
It's OK to have sex with someone on your first date	4.01	4.36	-0.35	0.694
It's OK to have sex with someone the same night you meet them	3.15	3.86	-0.71	0.375
It's OK to have sex with several different people in the same month	2.45	2.62	-0.17	0.808
It's okay to have sex without protection	2.46	2.12	0.34	0.630
It's OK to have sex with someone when you know they are someone else's girlfriend/boyfriend	2.79	2.86	-0.07	0.921
It's OK to have sex with someone if you are drunk or high	4.20	4.86	-0.66	0.467
It's OK to have sex with someone if you know they are drunk or high	2.77	2.62	0.16	0.819

Source: Follow-up survey administered 6 months after baseline.

Notes: Results in this table are based on 2,011 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c Score based on the seven items (shown below) represents the average percent of items agreed with by respondent for each group (alpha coefficient = 0.81).

* p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

4.4.3 Motivation

Students in both the treatment and control groups were highly motivated to avoid childbearing. There were no differences between the two groups on a composite measure or on any of three individual items. Youth in both groups indicated a strong belief in the importance of delaying childbearing until personal goals have been achieved (Exhibit 4.6).

There were no significant differences on the composite measure of motivation to avoid childbearing by replication site (Appendix Table A.5) or across subgroups.

Exhibit 4.6: Short-term Impacts on Motivation to Delay Childbearing

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value	Effect Size ^c
Motivation to delay childbearing ^d	3.69	3.69	0.00	0.907	-0.01
You have goals you want to accomplish before having a child	3.64	3.63	0.01	0.837	0.01
It is important for you to finish school before you have a child	3.72	3.74	-0.02	0.438	-0.04
It is important to have a job and a stable income before you have a child	3.72	3.71	0.01	0.795	0.01

Source: Follow-up survey administered 6 months after baseline.

Notes: Results in this table are based on 2,003 – 2,015 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c The effect size is the standardized effect size of the difference, which is the “treatment effect” divided by the pooled standard deviation of the treatment and control groups.

^d This scale averages responses to 3 items (shown in table) on attitudes toward childbearing and the importance of goal setting. Possible values range from 1 to 4 with higher values indicating greater motivation to delay childbearing (alpha coefficient = 0.87).

* p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

4.4.4 Intentions

¡Cuidate! did not affect student intentions to engage in sexual activity. Students in both treatment and control groups were equal in their expectation of engaging in sexual intercourse or oral sex in the 12 months after the survey. About 40 percent of youth in both groups reported that they intended to engage in these behaviors. Nearly all of the students reported their intention to use condoms or birth control if they were to engage in sexual intercourse (Exhibit 4.7).

There were no significant differences in program effects on intentions among sites (See Appendix Table A.6) or among any of the subgroups examined.

Exhibit 4.7: Short-term Impacts on Intentions

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value
Intentions				
Percentage of respondents reporting intentions to engage in the following behaviors in the next 12 months:				
Sexual intercourse	40.38	39.07	1.31	0.470
Oral sex	37.16	36.60	0.56	0.762
Use a condom if they were to have sexual intercourse	92.89	92.74	0.15	0.898
Use birth control if they were to have sexual intercourse	93.23	92.42	0.80	0.491

Source: Follow-up survey administered 6 months after baseline.

Notes: Results in this table are based on 1,996 – 2,005 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

* p< 0.05, ** p< 0.01, *** p< 0.001.

4.4.5 Skills

¡Cuidate! had a statistically significant impact on perceived condom negotiation skills but not on perceived refusal skills (Exhibit 4.8). Program participants were more likely to report that they could successfully negotiate condom use with a partner than were control group participants.

Exhibit 4.8: Short-Term Impacts on Skills

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value	Effect Size ^c
Skills (scale score)					
Perceived refusal skills ^d	3.19	3.13	0.06	0.062	0.08
Perceived condom negotiation skills ^e	3.53	3.46	0.07**	0.002	0.14

Source: Follow-up survey administered 6 months after baseline.

Notes: Results in this table are based on 2,015 – 2,016 respondents who provided valid survey responses to relevant items.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c The effect size is the standardized effect size of the difference, which is the “treatment effect” divided by the pooled standard deviation of the treatment and control groups.

^d This scale averages responses to six questions on perceived refusal skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills (alpha coefficient = 0.87).

^e This scale averages responses to seven questions on perceived condom skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills (alpha coefficient = 0.83).

* p< 0.05, ** p< 0.01, *** p< 0.001.

Site-Level Differences. Site-level analyses found a significant program impact on refusal skills in La Alianza. There, program participants reported significantly higher levels of refusal skills than their control group counterparts (See Appendix Table A.6).

Subgroup Differences. Program impacts on refusal skills varied by race/ethnicity. The impacts on refusal skills were greater for program participants who were of ‘Other’ race (See Appendix Table B.5).

4.5 Youth Sexual Behavior and Sexual Risk

Despite program impacts on youth knowledge, attitudes and one of the measures of skills, *¡Cuidate!* did not have a statistically significant impact on either of the two confirmatory outcome measures of current sexual activity or sexual intercourse without birth control (highlighted in Exhibit 4.9) at the short-term follow-up, 6 months post-baseline. Nor were there statistically significant differences in reported rates of other related sexual risk behaviors, although the differences on many of the outcomes favored the control group.

There were no significant effects on any behavioral outcomes in any of the three sites (See Appendix Table A.7).

Subgroup Differences. Although impact estimates for recent sexual behavior or risky sexual behavior were not statistically significant for the full sample, there were two sets of findings where impacts varied by subgroup and, within particular subgroups, unintended program effects were observed.

Sexual experience. There was variation in the impact on sexual intercourse in the last 90 days, depending on sexual experience at baseline. A significant unintended program effect was observed for those who had ever been sexually active at baseline, with treatment group members significantly more likely to report having had sexual intercourse in the last 90 days than their control group counterparts (See Appendix Table B.6). Among participants who had never engaged in sexual activity at baseline, those in the program group were no more likely than the control group to report engaging sexual intercourse in the last 90 days.

Race/Ethnicity. There was also variation in the impacts on oral sex and oral sex without a condom in the last 90 days, by respondent race/ethnicity. A significant and unintended program effect was observed for White program participants: treatment group members who were White were more likely to report having had oral sex and oral sex without a condom in the last 90 days than were their control group counterparts. There were no effects for Hispanic or Black participants, or participants who categorized themselves as “Other” race.

In the final chapter of the report, we consider these findings and their implications and explore the range of possible explanations.

Exhibit 4.9: Short-Term Impacts on Sexual Behavior and Sexual Risk

Outcome	Adjusted Treatment Mean ^a	Unadjusted Control Mean	Treatment Effect ^b	p-value
Sexual behavior (percentage responding affirmatively)				
Currently sexually active (in last 90 days) ^c	18.79	17.83	0.96	0.516
Sexual intercourse in the last 90 days	15.48	14.09	1.39	0.312
Oral sex in the last 90 days	14.69	13.13	1.56	0.266
Anal sex in the last 90 days	2.48	2.87	-0.39	0.704
Initiation of sexual activity ^d	14.62	12.86	1.76	0.303
Sexual risk (percentage responding affirmatively)				
Sexual intercourse without birth control (in last 90 days)	5.77	4.86	0.90	0.383
Sexual intercourse without a condom (in last 90 days)	9.81	8.10	1.70	0.157
Oral sex without a condom (in last 90 days)	12.93	11.25	1.68	0.211
Anal sex without a condom (in last 90 days) ^c	1.46	1.99	-0.53	0.525

Source: Follow-up survey administered 6 months after baseline.

Notes: Results are based on 2,009 - 2,012 respondents who provided valid survey responses to relevant items, except for the items measuring anal sex (n=1,173). Confirmatory outcomes are bolded.

^a The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^b The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates.

The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c Sexual activity is defined differently across grantees. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Touchstone.

^d The sample size for the initiation of sexual activity outcome is 1526, as this outcome only includes youth who were not sexually active at baseline.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests). (For the two confirmatory outcomes statistical significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$ implies statistical significance at these levels after applying a Benjamini-Hochberg adjustment for multiple comparisons.)

5. Discussion

This report on interim findings for the *¡Cuídate!* program model is the first of two and is not intended to provide comprehensive evidence about the most important behavioral outcomes—those that reflect the goals of the TPP initiative. A final assessment of the program’s effectiveness must await the findings from the longer-term follow-up survey, conducted 18 months after the program began. Nevertheless, these early results did not provide convincing evidence that *¡Cuídate!* reduced sexual risk-taking behaviors. Below, we provide insights into what was and was not accomplished by these ambitious efforts to replicate with fidelity a program that had been recognized as offering strong evidence of effectiveness (on the basis of the strength of the evaluation design).²²

¡Cuídate! was implemented with fidelity in all three replication sites and, in all three, succeeded in exposing youth to more information about sexual health and sexual risk prevention than they would otherwise have received. In general, in all three replication sites, students participated in enough of the program to meet the requirements imposed by OAH and the developer—in many cases, there was close to perfect attendance. In addition the program significantly increased students’ understanding of sexual risk behavior and ways to prevent outcomes like unplanned pregnancy and STIs. Student attitudes towards use of birth control and condoms were also significantly more positive as a result of program participation. Students in the group that received *¡Cuídate!* were also significantly more confident in their ability to negotiate the use of condoms.

Beyond these impacts, the program had no significant effects on motivation to delay pregnancy or intentions with respect to future sexual behavior. In addition, although there were significant overall program effects on knowledge, there was a cluster of items dealing with knowledge of STIs and their prevention for which half or fewer program participants knew the correct answer.

The program had no positive effect on any of the behavioral outcomes that represent the primary goals of this and all other TPP programs. To the contrary, there are findings that suggest the program may have had a negative impact on some of these outcomes for certain subgroups.

Given the lack of short-term impacts on key elements of the logic model that are hypothesized to influence behavioral outcomes (motivation and intentions) as well as the absence of positive impacts on sexual risk behaviors, we think it unlikely that the pattern of findings will change in any important way with the inclusion of longer-term impacts. For that reason, it seems useful to ask what could explain the lack of evidence for the program’s effectiveness. Typically, there are several potential explanations for lack of program effect: 1) poor implementation; 2) failure to reach the population for whom the program was intended; and 3) failure of the theory underlying the program. We can eliminate the first explanation immediately; the program was implemented with fidelity in all three replications and was successful in retaining participants.

The second explanation needs more examination. The program is specifically designed for Hispanic youth, broadly defined, although the developer maintains that its messages are appropriate for youth of all

²² The Teen Pregnancy Prevention Evidence Review (<http://tppevidencereview.aspe.hhs.gov/>) categorized the study of *¡Cuídate!*’ study design as “strong.”

ethnicities, since it is solidly based on another evidence-based program (EBP) that is intended for a broader population. All three replications successfully recruited Hispanic youth, although they also served non-Hispanic youth, since it is impossible in a school setting to target students of specific ethnic origin. Subgroup analyses that revealed differences between racial/ethnic groups, however, did not reveal any significant program impacts on Hispanic youth. Additional sensitivity analyses, in which we ran the full set of analyses on Hispanic youth only, also indicated similar findings to those of our full sample. The program model also does not specifically limit its likely effectiveness to low-income youth, since the rationale for its focus is a higher rate of HIV and other STIs among Hispanic youth in general (although the population on which it was originally tested was predominantly low-income). The TPP grant program, however, does give priority to low-income youth generally, since they are seen as at greater risk for adverse outcomes. Only one of the three replications met this criterion of serving primarily low-income youth consistently: La Alianza served a predominantly low-income youth population. While all three served populations that were believed to be at risk, given local statistics on pregnancy and STIs among youth, our data suggest that on other measures of risk, such as drug and alcohol use, and sexual experience and activity before the study began, the populations served in the three replication sites did not stand out as high risk after all.

The third explanation has two elements: a definition of the population for whom the developer believes the program will be effective; and the underlying theory of how the program achieves its goals. The first element also has two components: its ethnic focus and the age of youth for whom it is appropriate. The program is designated as appropriate for youth ages 13-19, and almost all of the study participants fell within this window. On average, youth in the study were between 14 and 15 years of age when they entered the study (and the program), and in one replication, which continues more than one-third of the study sample, all participants were at the low extreme of the age-range for which the program was designed. While still within the appropriate range, the fact that almost none were sexually active when they entered the program may affect the study's ability to detect effects on sexual risk behavior at this young age over the period covered by this report (i.e., six months). The program is designated as culturally sensitive to Hispanic youth generally, although communicating information and messages that are appropriate for a wider population. In reality, in the study that serves as the basis for *¡Cuidate!*'s designation as an EBP, the population served was Puerto Rican, a population that is distinct in many ways from other Hispanic groups. The Hispanic populations served in the three replications were more diverse: in the two Western sites, the populations served were almost entirely Mexican-American, while the Boston site served Puerto Rican, Central American and Dominican Hispanic populations. Pregnancy and STI risks are much higher for Puerto Rican youth than for Mexican-American and other Hispanic youth.

Aspects of the program model and its underlying theory of how it leads to the intended outcomes also warrant consideration. The program is designed to be delivered in six sessions and over a period of as little as two days and as long as six weeks. Possibly this is not sufficient time for the messages the program seeks to convey to be absorbed and translated into changes in intentions and actual behaviors. There may also be shortcomings in the program messages themselves and in the ways they are conveyed. Before the study began, grantees were concerned that the curriculum did not pay sufficient attention to pregnancy risk—understandable since the spur for its development was the rate of HIV and other STIs in Hispanic populations. For this reason, in two of the three sites, sessions were added to address this topic. Yet the program seemed to leave a substantial proportion of its participants unaware of important facts about STIs, even though it increased the proportion of youth who acquired correct knowledge relative to the control group. *¡Cuidate!* also seems to have been ineffective in changing the sexual behaviors of

youth that lead to STIs and their transmission, despite having shown some effectiveness on intermediate outcomes theorized to lead to behavior change.

It is important to consider these possible explanations for the lack of program effectiveness demonstrated by this study. However, because we cannot dismiss the possibility that the final findings may be different from these preliminary findings, it is premature to settle on one explanation over another.

The TPP Replication Study was designed to address important research and policy questions about the effectiveness of evidence-based programs, and what happens when they are taken to scale, replicated with different populations, and in different settings. The three program models were intentionally selected to maximize what would be learned about different strategies and begin to address identified gaps in the teen pregnancy prevention research. This report, part of a larger set of reports on replications of evidence-based program models, provides important information on the early effectiveness of *¡Cuidate!*. Early reports on the other two models will shed additional light on our understanding of different strategies for addressing youth risk behavior and promoting healthy choices for youth. The final reports will provide more comprehensive and conclusive evidence on the effectiveness of these models on sexual risk-taking behaviors and their consequences.

References

- American Academy of Pediatrics. (2013). Policy statement: Condom use by adolescents. doi: 10.1542/peds.2013-2831.
- CDC. (2012). *Estimated HIV incidence among adults and adolescents in the US 2007-2010. HIV Surveillance Report vol. 17*. Atlanta, GA: Author.
- CDC-P. (2004). *HIV surveillance report, 2003, vol.15*. Atlanta, GA: Author.
- Constantine, J., Player D., Silva, T., Hallgren, K., Grider, M., and Deke, J. (2009). *An evaluation of teachers trained through different routes to certification, final report* (NCEE 2009-4043). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Gleason, P., Clark, M., Tuttle, C. and Dwyer E. (2010). *The Evaluation of Charter School Impacts: Executive Summary* (NCEE 2010-4030). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Goesling, B., Colman, C., Trenholm, C., Terzian, M. & Moore, K. (2014). *Programs to reduce teen pregnancy, sexually transmitted infections, and associated risk behaviors: A systematic review*. *Journal of Adolescent Health*, 54(5), 499-507.
- Jemmott, J. B., Jemmott, L. S., Fong, G. T., & McCaffree, K. (1999). *Reducing HIV risk associated sexual behavior among African American adolescents: Testing the generality of intervention effects*. *American Journal of Community Psychology*, 27(2), 161–187.
- Jemmott III, J. B. (1992). *Reductions in HIV risk-associated sexual behaviors among black male adolescents: Effects of an AIDS prevention intervention*. *American Journal of Public Health*, 82(3), 372–377.
- Orr, L. L. (1999). *Social experiments: Evaluating public programs with experimental methods*. Thousand Oaks, CA: Sage Publications.
- 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.
- Schochet, P.Z. (2008a). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics*, 33(1), 62-87.
- Schochet P.Z. (2008b). Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations (NCEE 2008-4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- U.S. Department of Health and Human Services. (2015). HHS Teen Pregnancy Prevention Evidence Review [website]. <http://tppevidencereview.aspe.hhs.gov/>

REFERENCES

- U.S. Department of Health and Human Services. (2010). Teenage pregnancy prevention: Replication of evidence-based programs. Funding opportunity announcement and application instructions. Office of Adolescent Health, Office of Public Health and Science.
- U.S. Census Bureau. (2009). CB09-75. Washington, DC: Author.
- U.S. Census Bureau. (2012). CB12-FF19. Washington, DC: Author.
- Villaruel, A. M., Jemmott, J. B., & Jemmott, L. S. A randomized controlled trial testing an HIV prevention intervention for Latino youth. (2006). *Archives of Pediatrics & Adolescent Medicine*, 160(8), 772–777
- Villaruel A. M., Jemmott, L. S., & Jemmott, J. B. III. (2005). Designing a culturally-based intervention to reduce HIV sexual risk for Latino adolescents. *The Journal of Association of Nurses in AIDS Care*, 16, 23-31.

Appendix A: Site-Level Impacts

This study was carefully designed such that when data from all three replication sites were pooled into a single analysis, the combined sample would be large enough for the study to be adequately powered to detect effects of the *¡Cuidate!* intervention on all of the outcomes of interest. Although the pooled analysis is the primary focus of this study, there was clearly considerable interest on the part of study stakeholders in examining the results from each of the three replication sites, and the large sample sizes preserve the ability to conduct these analyses. Therefore this appendix presents site-specific impact estimates for each of the outcomes reported in the main text. We urge two major types of caution for readers who examine the results from the individual sites. The first is that the study was not designed to have large enough sample sizes in each individual site to have a good chance of detecting a treatment effect for all of the outcomes of interest. Thus, in a single site, lack of statistical significance could be the result of either an insufficiently large sample to detect a true effect, or it could mean that the intervention did not produce an effect on the outcome. Second, there are a large number of results presented in Appendix A, and these results are not adjusted for multiple comparisons. Some statistically significant findings would be expected purely by chance among such a large number of tests. Therefore, the findings in these tables should be interpreted with caution. The final column of each table shows the statistical result for a test of differences in the treatment effect across sites. When a statistically significant difference is found, the corresponding site-specific impacts are discussed in the main text, as we only interpret site-specific impacts when a significant difference among sites is found.

A.1 Impacts on Exposure to Program Information, by Site

Outcome	Community Action Partnership (n=490)				La Alianza (n=691)				Touchstone (n=840)				p-value for the Test of Differences Across Sites ^a
	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	
Percentage of respondents that reported receiving information on the following topics ^d :													
Relationships or marriage	92.42	86.21	6.21 *	0.046	89.19	80.63	8.56 **	0.002	87.58	80.29	7.29 **	0.003	0.846
Abstinence from sex	88.22	73.89	14.33 ***	0.000	84.52	70.24	14.28 ***	0.000	79.32	60.63	18.69 ***	0.000	0.509
Birth control methods	90.26	74.38	15.88 ***	0.000	84.00	70.24	13.76 ***	0.000	73.33	43.10	30.23 ***	0.000	0.000 ***
Where to obtain birth control	89.95	74.88	15.07 ***	0.000	80.45	65.61	14.84 ***	0.000	62.55	27.14	35.41 ***	0.000	0.000 ***
Sexually transmitted infections	92.88	78.71	14.17 ***	0.000	92.78	83.79	8.99 ***	0.001	90.38	65.62	24.76 ***	0.000	0.000 ***
How to talk with partner about sex and birth control	91.35	67.00	24.35 ***	0.000	86.54	70.24	16.30 ***	0.000	80.98	42.86	38.12 ***	0.000	0.000 ***
How to say no to sex	94.60	77.83	16.77 ***	0.000	87.85	77.29	10.56 ***	0.000	88.56	70.20	18.36 ***	0.000	0.102
How babies are made	92.96	88.18	4.78	0.070	92.37	87.30	5.07 *	0.028	91.61	87.14	4.47 *	0.031	0.981

Source: Follow-up survey administered six months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d Refers to information received in the 12 months prior to the survey administration.

p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

A.2 Impacts on Knowledge of Pregnancy Risk and STI Risk by Site

Outcome	Community Action Partnership (n=490)				La Alianza (n=692)				Touchstone (n=840)				p-value for the Test of Differences Across Sites ^a
	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	
Knowledge of pregnancy risk (percent of items respondents answered correctly) ^d	77.06	71.43	5.63 *	0.019	63.64	62.85	0.79	0.707	64.23	53.50	10.73 ***	0.000	0.002 **
Percentage of respondents correctly answering each item:													
Used correctly, how much can birth control pills reduce pregnancy risk?	73.98	67.49	6.49	0.132	58.61	58.10	0.51	0.892	55.63	44.86	10.77 **	0.001	0.125
Used correctly, how much can condoms reduce pregnancy risk?	77.14	67.98	9.16 *	0.031	62.51	63.24	-0.73	0.843	61.46	46.29	15.17 ***	0.000	0.006 **
A couple that has had unprotected sex and not gotten pregnant does not have to worry about getting pregnant.	85.57	85.71	-0.14	0.966	84.58	80.24	4.34	0.138	84.15	76.00	8.15 **	0.002	0.148
A woman is protected from pregnancy the day she begins taking the pill.	71.56	64.53	7.03	0.104	48.87	49.80	-0.93	0.805	55.58	46.86	8.72 **	0.010	0.141
Knowledge of STI risk (percent of items respondents answered correctly) ^e	69.59	62.11	7.48 ***	0.000	64.03	57.08	6.95 ***	0.000	60.38	44.79	15.59 ***	0.000	0.000 ***
Percentage of respondents correctly answering each item:													
Once you are infected with HIV you are infected for life	84.23	84.73	-0.50	0.894	81.83	74.70	7.13 *	0.030	81.35	56.29	25.06 ***	0.000	0.000 ***
There is a vaccine to prevent girls from getting HPV	56.46	45.32	11.14 *	0.013	49.45	49.01	0.44	0.910	31.79	36.57	-4.78	0.171	0.019 *
All STD/STIs can be cured by taking medicine	77.18	69.95	7.23	0.081	69.25	59.68	9.57 **	0.008	78.97	54.29	24.68 ***	0.000	0.001 ***
A person with an STD/STI who looks and feels healthy cannot transmit the infection to others	85.34	84.73	0.61	0.863	80.87	75.89	4.98	0.105	82.53	69.43	13.10 ***	0.000	0.014 *
Some STDs/STIs put you at greater risk of HIV	67.38	61.08	6.30	0.142	71.48	68.38	3.10	0.406	74.95	56.00	18.95 ***	0.000	0.004 **
About 1 out of 4 sexually active teens gets an STD/STI every year	72.03	67.49	4.54	0.306	64.71	56.52	8.19 *	0.034	66.82	56.86	9.96 **	0.004	0.628
You can get an STD/STI from having oral sex	80.05	73.40	6.65	0.099	77.87	68.38	9.49 **	0.007	67.20	50.86	16.34 ***	0.000	0.125
Used correctly, how much can condoms decrease the risk of HIV?	58.18	45.81	12.37 **	0.006	50.80	47.83	2.97	0.445	54.81	38.29	16.52 ***	0.000	0.032 *
You can get HIV even if you unprotected sex only 1 or 2 times	74.67	73.89	0.78	0.852	71.10	62.06	9.04 *	0.012	70.71	51.14	19.57 ***	0.000	0.001 **
Used correctly, how much can condoms decrease the risk of gonorrhea?	44.34	30.05	14.29 ***	0.000	36.26	32.81	3.45	0.331	28.01	22.57	5.44	0.089	0.110
Used correctly, how much can birth control pills decrease the risk of HIV?	68.12	55.17	12.95 **	0.002	59.46	44.66	14.80 ***	0.000	53.01	26.29	26.72 ***	0.000	0.009 **
Used correctly, how much can birth control pills decrease the risk of gonorrhea	66.85	53.69	13.16 ***	0.001	54.56	45.06	9.50 **	0.006	33.38	18.86	14.52 ***	0.000	0.547

Source: Follow-up survey administered 6 months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d Score based on the four items below. Values represent the average percent of items answered correctly by respondent for each group.

^e Score based on the 12 items below. Values shown represent the average percent of items answered correctly by respondent for each group.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

A.3 Impacts on Attitudes Toward Protection by Site

Outcome	Community Action Partnership (n=490)					La Alianza (n=692)					Touchstone (n=840)					p-value for the Test of Differences Across Sites ^a
	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	SES ^d	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	SES ^d	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	SES ^d	
Attitudes toward protection ^e	3.29	3.23	0.06	0.064	0.14	3.25	3.21	0.04	0.127	0.10	3.20	3.03	0.17 ***	0.000	0.42	0.001 ***
Birth control pills should always be used if a person your age has sexual intercourse	3.36	3.42	-0.06	0.383	-0.08	3.30	3.27	0.03	0.581	0.04	3.19	3.23	-0.04	0.491	-0.05	0.543
Birth control is too much trouble to use (reverse)	3.20	3.15	0.06	0.398	0.07	3.25	3.17	0.08	0.192	0.10	3.19	3.00	0.19 ***	0.000	0.25	0.212
Birth control is pretty easy to get	3.00	2.89	0.11	0.129	0.13	2.98	2.98	0.00	0.950	0.00	2.54	2.48	0.06	0.295	0.07	0.487
Birth control is important to make sex safer	3.39	3.34	0.05	0.403	0.07	3.28	3.29	-0.01	0.838	-0.02	3.19	3.17	0.01	0.769	0.02	0.746
Birth control has too many side effects (reverse)	2.82	2.65	0.17 **	0.006	0.23	2.69	2.63	0.06	0.276	0.08	2.76	2.55	0.22 ***	0.000	0.29	0.118
Using birth control is morally wrong (reverse)	3.39	3.29	0.10	0.080	0.15	3.31	3.28	0.04	0.490	0.05	3.30	3.13	0.17 ***	0.000	0.25	0.134
Condoms are too much trouble to use (reverse)	3.39	3.36	0.04	0.605	0.04	3.30	3.26	0.04	0.481	0.05	3.36	3.14	0.22 ***	0.000	0.28	0.036 *
Condoms are pretty easy to get	3.34	3.40	-0.06	0.379	-0.08	3.42	3.40	0.02	0.746	0.02	3.12	2.92	0.20 ***	0.000	0.25	0.006 **
Condoms are important to make sex safer	3.67	3.59	0.08	0.169	0.12	3.63	3.58	0.06	0.260	0.09	3.68	3.44	0.24 ***	0.000	0.37	0.011 *
Using condoms means you don't trust your partner (reverse)	3.49	3.38	0.11	0.088	0.15	3.46	3.45	0.01	0.821	0.02	3.48	3.26	0.23 ***	0.000	0.31	0.015 *
Using condoms is morally wrong (reverse)	3.56	3.47	0.09	0.112	0.14	3.55	3.51	0.04	0.463	0.06	3.60	3.28	0.32 ***	0.000	0.49	0.000 ***
Condoms decrease sexual pleasure (reverse)	2.78	2.78	0.00	0.968	0.00	2.85	2.68	0.17 **	0.009	0.20	2.95	2.69	0.26 ***	0.000	0.31	0.021 *

Source: Follow-up survey administered 6 months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d The effect size is the standardized effect size of the difference, which is the "treatment effect" divided by the pooled standard deviation of the treatment and control groups.

^e This construct averages responses to 12 items (shown in table) on attitudes towards condoms and birth control. Possible values range from 1 to 4 with higher values indicating more positive attitudes toward protection

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

A.4 Impacts on Attitudes Toward Risky Sexual Behavior by Site

Outcome	Community Action Partnership (n=490)				La Alianza (n=688)				Touchstone (n=833)				p-value for the Test of Differences Across Sites ^a
	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	
Support for risky sexual behavior (percent of items respondents agreed with) ^d	4.08	3.80	0.28	0.790	2.72	3.40	-0.68	0.455	2.87	3.00	-0.13	0.877	0.780
Percentage of respondents agreeing with each item:													
It's OK to have sex with someone on your first date	4.92	5.91	-0.99	0.569	2.94	5.16	-2.22	0.145	4.44	2.87	1.57	0.251	0.164
It's OK to have sex with someone the same night you meet them	5.24	4.93	0.31	0.844	2.27	3.17	-0.90	0.516	2.56	3.74	-1.18	0.341	0.747
It's OK to have sex with several different people in the same month	3.01	4.43	-1.42	0.303	3.00	2.38	0.62	0.607	1.68	1.72	-0.04	0.974	0.530
It's okay to have sex without protection	2.73	1.97	0.76	0.589	2.64	2.78	-0.14	0.909	2.20	1.72	0.48	0.667	0.879
It's OK to have sex with someone when you know they are someone else's girlfriend/boyfriend	3.19	2.46	0.73	0.625	2.81	3.17	-0.36	0.782	2.53	2.87	-0.34	0.771	0.822
It's OK to have sex with someone if you are drunk or high	6.44	5.42	1.02	0.572	2.75	4.37	-1.62	0.304	3.97	4.89	-0.92	0.515	0.529
It's OK to have sex with someone if you know they are drunk or high	3.02	1.48	1.54	0.263	2.65	2.78	-0.13	0.913	2.70	3.16	-0.46	0.672	0.498

Source: Follow-up survey administered 6 months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d Score based on the seven items (shown below) represents the average percent of items agreed with by respondent for each group.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

A.5 Impacts on Motivation to Delay Childbearing by Site

Outcome	Community Action Partnership (n=490)					La Alianza (n=689)					Touchstone (n=836)					p-value for the Test of Differences Across Sites ^a
	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	SES ^d	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	SES ^d	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p-value	SES ^d	
Motivation to delay childbearing ^e	3.66	3.69	-0.03	0.552	-0.05	3.71	3.74	-0.04	0.402	-0.07	3.70	3.66	0.04	0.299	0.07	0.345
You have goals you want to accomplish before having a child	3.59	3.62	-0.04	0.566	-0.05	3.66	3.68	-0.02	0.670	-0.03	3.66	3.61	0.06	0.247	0.08	0.404
It is important for you to finish school before you have a child	3.70	3.74	-0.04	0.463	-0.07	3.73	3.79	-0.06	0.228	-0.10	3.72	3.70	0.02	0.652	0.03	0.450
It is important to have a job and a stable income before you have a child	3.69	3.71	-0.02	0.718	-0.03	3.74	3.77	-0.02	0.612	-0.04	3.72	3.67	0.05	0.252	0.08	0.441

Source: Follow-up survey administered 6 months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d The effect size is the standardized effect size of the difference, which is the "treatment effect" divided by the pooled standard deviation of the treatment and control groups.

^e This scale averages responses to 3 items (shown in table) on attitudes toward childbearing and the importance of goal setting. Possible values range from 1 to 4 with higher values indicating greater motivation to delay childbearing.

* p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

A.6 Impacts on Intentions and Skills by Site

Outcome	Community Action Partnership (n=489)					La Alianza (n=691)					Touchstone (n=837)					p-value for the Test of Differences Across Sites ^a
	Adj. T Mean	Unadj. C Mean	T Effect ^b	p-value	SES ^c	Adj. T Mean	Unadj. C Mean	T Effect ^b	p-value	SES ^c	Adj. T Mean	Unadj. C Mean	T Effect ^b	p-value	SES ^c	
Intentions																
Percentage of respondents reporting intentions to engage in the following behaviors in the next 12 months:																
Sexual intercourse	50.71	45.32	5.39	0.134		52.58	52.40	0.18	0.954		25.36	25.66	-0.30	0.917		0.419
Oral sex	46.57	45.05	1.52	0.678		45.81	44.40	1.41	0.658		25.23	25.95	-0.72	0.802		0.843
Use a condom if they were to have sexual intercourse	90.60	93.07	-2.47	0.299		92.06	90.00	2.06	0.322		94.75	94.52	0.23	0.903		0.358
Use birth control if they were to have sexual intercourse	94.81	92.54	2.27	0.326		91.22	92.37	-1.15	0.567		93.88	92.40	1.48	0.413		0.476
Skills																
Perceived refusal skills (scale score) ^e	3.12	3.20	-0.08	0.231	-0.10	3.25	3.10	0.15 **	0.007	0.21	3.18	3.11	0.07	0.157	0.10	0.026 *
Perceived condom negotiation skills (scale score) ^f	3.51	3.51	0.00	0.922	0.01	3.58	3.50	0.09 *	0.029	0.17	3.50	3.41	0.10 **	0.007	0.19	0.248

Source: Follow-up survey administered 6 months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. For outcomes reported as percentages, the treatment effect is expressed as a difference in percentage points. For scale outcomes, the treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d The effect size is the standardized effect size of the difference, which is the "treatment effect" divided by the pooled standard deviation of the treatment and control groups.

^e This scale averages responses to 6 questions on perceived refusal skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills.

^f This scale averages responses to 7 questions on perceived condom skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills.

* p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

A.7 Impacts on Sexual Behavior and Sexual Risk by Site

Outcome	Community Action Partnership (n=486)				La Alianza (n=688)				Touchstone (n=838)				p-value for the Test of Differences Across Sites ^a
	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	Adj. T Mean ^b	Unadj. C Mean	T Effect ^c	p- value	
Sexual behavior (percentage responding affirmatively)													
Currently sexually active in the last 90 days ^d	26.74	24.26	2.48	0.398	28.61	28.97	-0.36	0.889	7.12	6.03	1.09	0.636	0.764
Sexual intercourse in the last 90 days	21.65	17.82	3.83	0.161	24.20	24.21	-0.01	0.996	5.63	4.60	1.03	0.630	0.556
Oral sex in the last 90 days	23.24	19.31	3.93	0.159	20.90	20.72	0.18	0.941	5.25	4.03	1.22	0.577	0.587
Anal sex in the last 90 days	0.69	1.98	-1.29	0.407	3.89	3.59	0.30	0.826					0.441
Sexual risk (percentage responding affirmatively)													
Sexual intercourse without birth control in the last 90 days	8.81	6.44	2.37	0.250	8.68	8.33	0.35	0.846	1.89	1.44	0.45	0.777	0.712
Sexual intercourse without a condom in the last 90 days	13.43	8.91	4.52	0.058	16.49	15.08	1.41	0.497	2.80	2.59	0.21	0.909	0.358
Oral sex without a condom in the last 90 days	21.48	16.34	5.14	0.053	18.05	17.93	0.12	0.957	4.25	3.46	0.79	0.703	0.313
Anal sex without a condom in the last 90 days ^d	0.39	1.98	-1.59	0.210	2.27	1.99	0.28	0.803					0.267

Source: Follow-up survey administered 6 months after baseline.

^a This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

^b The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

^c The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^d Sexual activity is defined differently across grantees. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Touchstone.

* p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests).

Appendix B: Subgroup Impacts

B.1 Impacts on Exposure to Program Information at Interim Follow-up, by Subgroup

	Treatment Effect ^a	p-value ^b
Learned about birth control methods		
Subgroup: Respondent age		
Respondent less than age 15 (n=1,141)	26.25***	0.000
Respondent age 15 or older (n=876)	14.39***	0.000
Learned about where to obtain birth control		
Subgroup: Respondent age		
Respondent less than age 15 (n=1,143)	28.67***	0.000
Respondent age 15 or older (n=877)	16.53***	0.000
Learned about STIs		
Subgroup: Respondent age		
Respondent less than age 15 (n=1,143)	20.17***	0.000
Respondent age 15 or older (n=876)	12.45***	0.000
Learned about talking to partner about having sex or using birth control		
Subgroup: Respondent age		
Respondent less than age 15 (n=1,142)	34.18***	0.000
Respondent age 15 or older (n=878)	18.47***	0.000
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=1,530)	29.73***	0.000
Ever sexually active at baseline (n=490)	19.60***	0.000
Learned about how to say no to sex		
Subgroup: Respondent gender		
Male (n=902)	21.53***	0.000
Female (n=1,117)	10.35***	0.000
Learned about relationships or marriage		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=1,531)	4.52*	0.011
Ever sexually active at baseline (n=490)	17.05***	0.000
Learned about how babies are made		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=1,531)	3.22*	0.034
Ever sexually active at baseline (n=490)	9.77***	0.000

Source: Follow-up survey administered 6 months after baseline.

Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < 0.05$). For example, a test result indicated that the treatment effect on learning about STIs was significantly different for younger and older respondents.

APPENDIX B: SUBGROUP IMPACTS

^a This column shows the estimated treatment effect (Treatment/control difference in the percent reporting receiving information) for the subgroup indicated in the row

^b This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

B.2 Impacts on Knowledge of Pregnancy Risk, by Subgroup

	Treatment Effect ^a	p-value ^b
Subgroup: Respondent age		
Respondent less than age 15 (n=1,144)	8.37***	0.000
Respondent age 15 or older (n=878)	3.19	0.083
Subgroup: Respondent gender		
Male (n=903)	3.13	0.083
Female (n=1,119)	8.56***	0.000

Source: Follow-up survey administered 6 months after baseline.

Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < 0.05$). For example, a test result indicated that the treatment effect on knowledge of pregnancy risk was significantly different for younger respondents.

^a This column shows the estimated treatment effect (Treatment/control difference in the average percent of items answered correctly) for the subgroup indicated in the row.

^b This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

B.3 Impacts on Knowledge of STI Risk, by Subgroup

	Treatment Effect ^a	p-value ^b
Subgroup: Respondent age		
Respondent less than age 15 (n=1,144)	14.17***	0.000
Respondent age 15 or older (n=878)	6.05***	0.000
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=1,532)	12.07***	0.000
Ever sexually active at Baseline (n=490)	5.99**	0.001

Source: Follow-up survey administered 6 months after baseline.

Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < 0.05$). For example, a test result indicated that the treatment effect on knowledge of STI risk was significantly different for younger versus older respondents.

^a This column shows the estimated treatment effect (treatment/control difference in the average percent of items answered correctly) for the subgroup indicated in the row.

^b This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

APPENDIX B: SUBGROUP IMPACTS

B.4 Impacts on Attitudes toward Protection^a, by Subgroup

	Treatment Effect ^b	p-value ^c
Subgroup: Respondent age		
Respondent less than age 15 (n=1,144)	0.15***	0.000
Respondent age 15 or older (n=878)	0.04	0.112

Source: Follow-up survey administered 6 months after baseline.

Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < 0.05$). For example, a test result indicated that the treatment effect on attitudes toward protection was significantly different for younger versus older respondents.

^a Possible values range from 1 to 4 with higher values indicating more positive attitudes towards protection.

^b This column shows the estimated treatment effect (Treatment/control difference) for the subgroup indicated in the row.

^c This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

B.5 Impacts on Refusal Skills^a, by Subgroup

	Treatment Effect ^b	p-value ^c
Subgroup: Respondent race/ethnicity		
Hispanic (n=1,433)	0.07	0.055
Black (n=96)	0.23	0.107
White (n=359)	-0.14	0.065
Other (n=127)	0.37**	0.005

Source: Follow-up survey administered 6 months after baseline.

Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < 0.05$). For example, a test result indicated that the treatment effect on refusal skills was significantly different across the replication sites.

^a Possible values range from 1 to 4 with higher values indicating greater perceived skills.

^b This column shows the estimated treatment effect (Treatment/control difference) for the subgroup indicated in the row.

^c This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

B.6 Impacts on Sexual Risk Behavior, by Subgroup

	Treatment Effect ^a	p-value ^b
Sexual Intercourse in the Last 90 Days		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=1,526)	-0.40	0.798
Ever sexually active at baseline (n=486)	7.41*	0.010
Oral Sex in the Last 90 Days		
Subgroup: Respondent race/ethnicity		
Hispanic (n=1,430)	-0.78	0.640
Black (n=96)	-1.10	0.860
White (n=357)	9.15**	0.005
Other (n=126)	7.90	0.164
Oral Sex Without a Condom in the Last 90 Days		
Subgroup: Respondent race/ethnicity		
Hispanic (n=1,430)	-0.77	0.631
Black (n=96)	-0.18	0.975
White (n=357)	8.69**	0.006
Other (n=126)	10.12	0.061

Source: Follow-up survey administered 6 months after baseline.

Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < 0.05$). For example, a test result indicated that the treatment effect on current oral sex was significantly different among the racial-ethnic groups.

^a This column shows the estimated treatment effect (Treatment/control difference in percent reporting engaging in the behavior) for the subgroup indicated in the row.

^b This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Appendix C: Measures

The measures we used to examine short-term program impacts stem from our research questions (Section 3.1) and logic model (Exhibit 2.2) and are organized into three categories:

- Exposure to information;
- Intermediate outcomes; and
- Youth sexual behavior.

Measures in the first category (*exposure to information*) reflect receipt of sexuality education and reproductive health information. These provide insight into *¡Cuidate!*'s success in reaching youth. Measures of *intermediate outcomes* indicate the extent to which youth assimilated the program's messages and reflected them in their knowledge, attitudes, motivation, intentions, and skills—all of which are hypothesized precursors of change in youth's sexual behavior. Measures of *youth sexual behavior* include recent sexual activity, and recent sexual risk behavior (unprotected sexual activity). In the sections that follow, we describe each category by defining constituent measures and their construction.

C.1 Exposure to Information

To assess whether *¡Cuidate!* increased exposure to information on sexual health, contraception, and STI transmission and prevention, at the short-term follow-up, we asked youth about their receipt of sexuality education and reproductive health information.²³ On the survey, they responded to a series of questions asking about their exposure to information about: (a) relationships or marriage; (b) abstinence from sex; (c) birth control methods; (d) where to obtain birth control; (e) STIs; (f) how to talk with a partner about sex and birth control; (g) how to say no to sex; and (h) how babies are made. For each, youth were asked whether they had “received information or learned about” the topic in the 12 months prior to survey administration. Responses were coded in a binary fashion, as 1 = “yes” and 0 = “no.”

C.2 Intermediate Outcomes

Intermediate outcomes are those expected to portend changes in behavior. At the short-term follow-up, we asked youth a wide variety of questions to gauge their understanding, thoughts, beliefs, and perceptions of topics addressed by the program. We organized these measures conceptually into five domains: knowledge, attitudes, motivation, intentions, and skills. Using survey items relevant to each domain, we conducted principal component factor analyses and reliability testing to construct composite measures in each domain, where this was possible. In addition, we used baseline data (when the same items were asked) to examine the stability over time of composite measures, and examined the follow-up data by racial-ethnic subgroup to assess the stability of constructs.

²³ At baseline and before random assignment, youth were asked these same questions about the 12 month period preceding the study. The reference period for the follow-up survey included the period in which treatment group members were offered the intervention (and controls were not) as well as the six months prior to random assignment.

Knowledge

To examine program-related changes in youth's sexual health knowledge, we constructed two measures: *knowledge of pregnancy risk* and *knowledge of STI risk*. These measures were defined conceptually and constructed to differentiate accurate knowledge from misinformation. They may be considered tests of understanding of the factors contributing to pregnancy and STIs. The construction of these measures is described below and detailed information about their component items is presented in Exhibit C2.1.

- ***Knowledge of pregnancy risk*** is a composite measure that is the mean (multiplied by 100) of four binary variables regarding knowledge of the extent to which contraceptive methods can prevent pregnancy and circumstances under which pregnancy is possible (See Exhibit C2.1 for coding and other details). Scores on this scale range from 0 to 100 and represent the percentage of correct answers across the four items. Higher values indicate more accurate knowledge.
- ***Knowledge of STI risk*** is a composite measure that is the mean of 12 binary variables pertaining to knowledge of STI prevention, transmission, and treatment (see Exhibit C2.1 for coding and other details). Scores on this scale range from 0 to 100 and represent the percentage of correct answers across the 12 items. Higher values indicate more accurate knowledge.

Exhibit C2.1: Knowledge Scales and Component Items

Component Items	Coding
Knowledge of Pregnancy Risk (4 items)	
Used correctly, how much can birth control pills reduce pregnancy risk?	Youth responded to this question on a scale from 1= "Not at all" to 4="Completely." This item was recoded into a binary variable where the correct response ("a lot") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can condoms reduce pregnancy risk?	Youth responded to this question on a scale from 1= "Not at all" to 4="Completely." This item was recoded into a binary variable where the correct response ("a lot") was coded as 1 and all other responses were coded as 0.
A couple that has had unprotected sex and not gotten pregnant does not have to worry about getting pregnant.	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false, and 0 indicates they were sure or thought the statement was true or did not know.
A woman is protected from pregnancy the day she begins taking the pill.	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false and 0 indicates they were sure or thought the statement was true or did not know.
Knowledge of STI Risk (12 items)	
You can't get infected with HIV if you have sex only once or twice w/o a condom	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false and 0 indicates they were sure or thought the statement was true or did not know
Once you are infected with HIV you are infected for life	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was true and 0 indicates they were sure or thought the statement was false or did not know.
There is a vaccine to prevent girls from getting HPV	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was true and 0 indicates they were sure or thought the statement was false or did not know.
All STDs/STIs can be cured by taking medicine	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false and 0 indicates they were sure or thought the statement was true or did not know.
A person with an STD/STI who looks and feels healthy cannot transmit the infection to others	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false and 0 indicates they were sure or thought the statement was true or did not know.
Some STDs/STIs put you at greater risk of HIV	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was true and 0 indicates they were sure or thought the statement was false or did not know.
About 1 out of 4 sexually active teens gets an STD/STI every year	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was true and 0 indicates they were sure or thought the statement was false or did not know.

Component Items	Coding
You can get an STD/STI from having oral sex	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was true and 0 indicates they were sure or thought the statement was false or did not know.
Used correctly, how much can condoms decrease the risk of HIV?	Youth responded to this question on a scale from 1 = "Not at all" to 4 = "Completely." This item was recoded into a binary variable where the correct response ("a lot") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can condoms decrease the risk of gonorrhea?	Youth responded to this question on a scale from 1 = "Not at all" to 4 = "Completely." This item was recoded into a binary variable where the correct response ("a lot") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can birth control pills decrease the risk of HIV?	Youth responded to this question on a scale from 1 = "Not at all" to 4 = "Completely." This item was recoded into a binary variable where the correct response ("not at all") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can birth control pills decrease the risk of gonorrhea?	Youth responded to this question on a scale from 1 = "Not at all" to 4 = "Completely." This item was recoded into a binary variable where the correct response ("not at all") was coded as 1 and all other responses were coded as 0.

Attitudes

The short-term survey included 24 items querying attitudes toward sexual behaviors, sexual risks, and contraceptive methods. From among these, we constructed two measures to examine program impacts on youths’ sexual health attitudes: *attitudes toward protection* and *attitudes toward risky sexual behavior*. These measures are described below and detailed information about their component items is presented in Exhibit C2.2.

- ***Attitudes toward protection*** is a composite measure that is the mean of responses to 12 items about the importance of using condoms and/or birth control during sexual activity. (See Exhibit C2.2 for coding and other details.) Scores on this scale represent the level of support for using protection. They range from 1 to 4 with high scores indicating positive and supportive attitudes toward contraceptive use to prevent STIs and/or pregnancy. The measure demonstrates acceptable internal consistency reliability ($\alpha = 0.79$).²⁴
- ***Attitudes toward risky sexual behavior*** is a composite measure that is the mean of seven binary items (multiplied by 100) querying the acceptability and normativeness of risky sexual behaviors. (See Exhibit C2.2 for coding and other details.) Scores on this scale range from 0 to 100 and represent the percent of items agreed with: Higher values reflect more support for risky behavior. The measure demonstrates good internal consistency reliability ($\alpha = 0.81$).

²⁴ As a general rule of thumb, the internal consistency of scales with reliability coefficients between 0.70–0.79 is considered “acceptable,” between 0.80 – 0.89 is considered “good,” and 0.90 or greater is considered “excellent.”

Exhibit C2.2: Attitudes Scales and Component Items

Component Items	Coding
Attitudes Toward Protection (12 items)	
Birth control pills should always be used if a person your age has sexual intercourse.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward birth control.
Birth control is too much trouble to use.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward birth control.
Birth control is pretty easy to get.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward birth control.
Birth control is important to make sex safer.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward birth control.
Birth control has too many side effects.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward birth control.
Using birth control is morally wrong.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward birth control.
Condoms are too much trouble to use.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward condoms.
Condoms are pretty easy to get.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward condoms.
Condoms are important to make sex safer.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward condoms.
Using condoms means you don't trust your partner.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward condoms.
Using condoms is morally wrong.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward condoms.
Condoms decrease sexual pleasure.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4 = "Strongly disagree." High values indicate more positive attitudes toward condoms.
Attitudes Toward Risky Sexual Behavior (7 items)	
It's OK to have sex with someone on your first date.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.
It's OK to have sex with someone the same night you meet them.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.

APPENDIX C: MEASURES

Component Items	Coding
It's OK to have sex with several different people in the same month.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.
It's OK to have sex without protection.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.
It's OK to have sex with someone when you know they are someone else's girlfriend/boyfriend.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.
It's OK to have sex with someone if you are drunk or high.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.
It's OK to have sex with someone if you know they are drunk or high.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.

Motivation

The short-term survey included 22 items related to youth's motivation to engage in safe sexual practices and reduce their risk. From these, we developed a measure of motivation to delay childbearing. It is the average of three items related to reasons for avoiding childbearing. (See Exhibit C2.3 for coding and other details). Scores on this scale range from 1 to 4 with higher scores indicating more motivation to wait to have a child. The scale demonstrated good internal consistency reliability ($\alpha = 0.87$).

Exhibit C2.3: Motivation Scale and Component Items

Component Items	Coding
Motivation to Delay Childbearing (3 items)	
You have goals you want to accomplish before having a child.	Youth responded to this question on a scale from 1 = "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more agreement.
It is important for you to finish school before you have a child.	Youth responded to this question on a scale from 1 = "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more agreement.
It is important to have a job and a stable income before you have a child.	Youth responded to this question on a scale from 1 = "Strongly agree" to 4 = "Strongly disagree." We reverse coded this item so that higher values indicate more agreement.

Intentions

We used the four items presented in Exhibit C2.4 to examine impacts on youth's intended or anticipated sexual behavior in the coming year.

Exhibit C2.4: Intentions Measures

Item	Coding
Do you intend to have sexual intercourse in the next year, if you have the chance?	Youth responded to this question on a scale from 1 = "Yes, definitely" to 4 = "No, definitely not." This item was recoded into a binary variable where affirmative responses (definitely, probably) were coded as 1 and negative responses (definitely not, probably not) were coded as 0.
Do you intend to have oral sex in the next year, if you have the chance?	Youth responded to this question on a scale from 1 = "Yes, definitely" to 4 = "No, definitely not." This item was recoded into a binary variable where affirmative responses (definitely, probably) were coded as 1 and negative responses (definitely not, probably not) were coded as 0.
If you have sexual intercourse in the next year, do you intend to use birth control?	Youth responded to this question on a scale from 1 = "Yes, definitely" to 4 = "No, definitely not." This item was recoded into a binary variable where affirmative responses (definitely, probably) were coded as 1 and negative responses (definitely not, probably not) were coded as 0.
If you have sexual intercourse in the next year, do you intend to use a condom?	Youth responded to this question on a scale from 1 = "Yes, definitely" to 4 = "No, definitely not." This item was recoded into a binary variable where affirmative responses (definitely, probably) were coded as 1 and negative responses (definitely not, probably not) were coded as 0.

Skills

The short-term follow-up survey included items regarding skills important to reproductive health. We constructed two measures to examine program impacts on youth's perceived ability to say no to sex (*refusal skills*), and successfully negotiate condom use with a partner (*condom negotiation skills*). These measures are described below and detailed information about their component items is presented in Exhibit C2.5.

- ***Refusal skills*** is a composite measure that is the mean of responses to six items about perceived ability to say no to sex in a variety of situations. (See Exhibit C2.5 for coding and other details.) Scores on this scale range from 1 to 4 with high scores indicating more confidence in one's abilities to abstain from intercourse. The measure demonstrates good internal consistency reliability ($\alpha = 0.87$).
- ***Condom negotiation skills*** is a composite measure that is the mean of responses to seven items about perceived ability to obtain and negotiate the use of condoms. (See Exhibit C2.5 for coding and other details.) Scores on this scale range from 1 to 4 with high scores indicating more confidence in one's abilities to obtain and negotiate the use of condoms. The measure demonstrates good internal consistency reliability ($\alpha = 0.83$).

Exhibit C2.5: Skills Scales and Component Items

Component Items	Coding
Refusal Skills (6 items)	
How sure are you that you would be able to say no to having sexual intercourse if your partner really wanted to, but you were not ready?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
How sure are you that you would be able to say no to having sexual intercourse if you just met someone you really liked and that person wanted to have sex, but you didn't?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
How sure are you that you would be able to say no to having sexual intercourse if you had strong sexual feelings for that person?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
How sure are you that you would be able to say no to having sexual intercourse if neither you nor your partner had any form of birth control?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
How sure are you that you would be able to say no to having sexual intercourse if you have dated for a long time?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
How sure are you that you would be able to say no to having sexual intercourse after you have been drinking alcohol?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
Condom Negotiation Skills (7 items)	
If you were going to have sex could you get or buy a condom?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you talk about using condoms with your partner before having sex?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you insist on using a condom if your partner didn't want to use one?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you ask your partner to use condoms even if the two of you had sex before w/o using condoms?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you use a condom without spoiling the mood?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you ask a new partner to use condoms?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you get a partner to use condoms, even if you're drunk or high?	Youth responded to this question on a scale from 1 = "I'm sure I could" to 4 = "I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.

C.3 Youth Sexual Behavior and Sexual Risk

To understand program effects on youths’ sexual behavior and sexual risk, we examined their responses to questions about their history of sexual activity, their recent sexual behavior, and their recent sexual risk behavior. We used the nine items presented in Exhibit C.3.1 to examine impacts on sexual behavior and sexual risk.

Exhibit C.3.1: Youth Sexual Behavior and Sexual Risk Measures

Measure	Item	Coding
Sexual Behavior		
Initiation of sexual activity	Have you ever had any of the following: sexual intercourse, oral sex or anal sex?	Youth who were not sexually active at baseline responded to this question with a yes (1)/no(0) answer. This item was coded 0 or 1, with 1 representing one or more forms of sexual activity (sexual intercourse, oral sex, and/or anal sex) since baseline and 0 representing no sexual activity since baseline. Responses to other sexual behavior and sexual risk questions were examined and back-coded into this question such that youth who reported they had engaged in one or more of the sexual activities received a score of 1. Note that sexual activity is defined differently across grantees. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and/or anal sex. Youth were not asked about anal sex in Touchstone.
Currently sexually active (in last 90 days)	Coded from three separate items measuring sexual intercourse in the last 90 days, oral sex in the last 90 days, and anal sex in the last 90 days.	Youth who reported they had engaged in one or more of the sexual activities (sexual intercourse, oral sex, or anal sex) during the last 90 days received a score of 1 on this measure. Youth who reported no sexual activity during the last 90 days received a score of 0, as did those who reported (on a separate question) that they had never been sexually active. Note that sexual activity is defined differently across grantees. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Touchstone.
Sexual intercourse in the last 90 days	Now please think about the past 3 months. In the past 3 months, have you had sexual intercourse?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in sexual intercourse in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in sexual intercourse in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.
Oral sex in the last 90 days	Now please think about the past 3 months. In the past 3 months, have you had oral sex?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in oral sex in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in oral sex in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.
Anal sex in the last 90 days	Now please think about the past 3 months. In the past 3 months, have you had anal sex?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in anal sex in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in anal sex in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.

APPENDIX C: MEASURES

Measure	Item	Coding
Sexual Risk		
Sexual intercourse without birth control (in last 90 days)	In the past 3 months, have you had sexual intercourse without you or your partner using any of these methods of birth control, even just once? <ul style="list-style-type: none"> • Condoms • Birth control pills • The shot (Depo-Provera) • The patch • The ring (NuvaRing) • IUD (Mirena or Paragard) • Implants (Implanon) 	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in sexual intercourse without birth control in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in sexual intercourse without birth control in the last 90 days received a score of 0 on the measure, as did those who reported (on separate questions) that they had not had oral sex in the last 90 days or that they had never been sexually active.
Sexual intercourse without a condom (in last 90 days)	In the past 3 months, have you had sexual intercourse without you or your partner using a condom?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in sexual intercourse without a condom in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in sexual intercourse without a condom in the last 90 days received a score of 0 on the measure, as did those who reported (on separate questions) that they had not had sexual intercourse in the last 90 days or that they had never been sexually active.
Oral sex without a condom (in last 90 days)	In the past 3 months, have you had oral sex without using a condom, even once?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in oral sex without a condom in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in oral sex without a condom in the last 90 days received a score of 0 on the measure, as did those who reported (on separate questions) that they had not had oral sex in the last 90 days or that they had never been sexually active.
Anal sex without a condom (in last 90 days)	In the past 3 months, have you had anal sex without using a condom, even once?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in anal sex without a condom in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in anal sex without a condom in the last 90 days received a score of 0 on the measure, as did those who reported (on separate questions) that they had not had anal sex in the last days or that they had never been sexually active.

Appendix D: Supporting Tables

Exhibit D.1: Characteristics of the Analytic Sample at Baseline

Measure	Treatment Mean ^a	Control Mean	Group Difference ^b	p-value
Demographic characteristics				
Age				
Mean	14.39	14.37	0.02	0.629
Grade				
Mean	9.19	9.13	0.05*	0.015
Gender				
Female ^c	52.98	52.98	0.00	n/a
Race/ethnicity^d				
Hispanic	71.40	70.10	1.30	0.506
Black	3.73	5.83	-2.10*	0.027
White	17.96	18.49	-0.53	0.749
Other	6.91	5.58	1.32	0.231
Family structure and relationships				
Lives with biological parents	94.30	90.71	3.60**	0.003
Feels very close to and cared for by father	44.45	46.06	-1.61	0.499
Feels very close to and cared for by mother	59.16	60.29	-1.16	0.611
Risk behaviors				
Ever smoked cigarettes	18.64	18.41	0.23	0.897
Ever drank alcohol	46.06	48.08	-2.03	0.355
Ever used marijuana	25.32	25.90	-0.58	0.763
Knowledge				
Knowledge of pregnancy risk ^e	48.37	47.98	0.39	0.831
Knowledge of STI risk ^e	38.61	39.36	-0.75	0.556
Attitudes				
Attitudes toward protection ^f	3.07	3.06	0.00	0.830
Intentions				
Intentions to have oral sex in the next 12 months	25.10	22.76	2.34	0.205
Intentions to have sexual intercourse in the next 12 months	31.91	27.50	4.41*	0.023
Intentions to use a condom if they were to have sexual intercourse	92.99	94.72	-1.73	0.132
Intentions to use birth control if they were to have sexual intercourse	92.07	91.99	0.08	0.953

APPENDIX D: SUPPORTING TABLES

Measure	Treatment Mean ^a	Control Mean	Group Difference ^b	p-value
Sexual Behavior				
Ever sexually active ^g	25.22	21.99	3.23	0.074
Currently sexually active (in last 90 days)	17.38	14.69	2.69	0.095
Sexual intercourse in the last 90 days	14.49	12.37	2.12	0.164
Oral sex in the last 90 days	12.66	10.34	2.32	0.107
Anal sex in the last 90 days	4.25	3.42	0.84	0.484
Sexual Risk				
Sexual intercourse without a condom in the last 90 days	8.33	6.44	1.89	0.112
Oral sex without a condom in the last 90 days	11.31	9.17	2.14	0.119
Anal sex without a condom in the last 90 days	2.77	2.05	0.72	0.464
Sexual intercourse without birth control in the last 90 days	4.38	3.48	0.90	0.328
Baseline exposure to program information^h				
Relationships or marriage	78.32	78.03	0.29	0.877
Abstinence from sex	59.21	58.74	0.47	0.824
Birth control methods	53.18	51.40	1.78	0.393
Where to obtain birth control	44.63	44.94	-0.31	0.879
Sexually transmitted infections	74.00	72.08	1.92	0.296
How to talk with partner about sex and birth control	43.68	44.16	-0.48	0.821
How to say no to sex	65.63	66.20	-0.57	0.790
How babies are made	84.26	85.41	-1.14	0.479

Source: Baseline survey administered prior to randomization.

Notes: Results in this table are based on the analytic sample of 1,939 - 2,022 respondents who provided valid survey responses to relevant items except for the items measuring how close the respondent feels to their mother (n= 1,930) father (n= 1,800), intentions to use birth control (n= 1,923), and anal sex (1,143). Values shown are percentages unless otherwise indicated. The items that compose measures of attitudes toward risky sexual behavior, motivation to delay childbearing, refusal skills, and condom negotiation skills were not asked at baseline.

^a The treatment mean was calculated as the sum of the control group mean and the model estimated treatment-control difference (group difference).

^b The baseline treatment-control difference was estimated where the dependent variable was the baseline measure, and the only independent variables included in the model were the treatment group indicator and terms for the randomization blocks. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

^c The analytic model for outcomes estimates impacts within gender groups, and aggregates impacts across the groups. This approach induces exact baseline equivalence of treatment and control groups on gender.

^d Racial ethnic categories are Hispanic, black non-Hispanic, white non-Hispanic, and other race non-Hispanic, where other is defined as Asian, American Indian or Alaska native, native Hawaiian or other Pacific Islander, multiracial, or undisclosed race.

^e Knowledge variables are composite scale scores representing the proportion of items answered correctly.

^f Attitude variable is a composite scale score with higher scores indicating more positive attitudes.

^g Sexual activity is defined differently across grantees. In La Alianza and Community Action Partnership, sexual activity refers to sexual intercourse, oral sex, and/or anal sex. Youth were not asked about anal sex in Touchstone.

^h Refers to information received in the 12 months prior to the survey administration.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).