Identifying Programs That Impact Teen Pregnancy, Sexually Transmitted Infections, and Associated Sexual Risk Behaviors

Review Protocol

Version 1.0

Review Authors

Mathematica Policy Research
Child Trends
Contact: pprer@mathematica-mpr.com

Review Timeline

Initial review: September 2009 – April 2010
First update: December 2010 – ongoing

TABLE OF CONTENTS

BACKGROUND

The consequences of adolescent sexual activity remain a troubling issue in the U.S. Nationwide, 46 percent of high school students have had sexual intercourse, and nearly 21 percent report having had four or more partners by graduation [1]. In 2009, nearly 39 percent of sexually active high school students had not used a condom during their last sexual intercourse [1]. These behaviors increase the risks of pregnancy and sexually transmitted infections (STIs). In 2009, there were approximately 39 births per 1,000 females 15 to 19 years of age [2]. Estimates suggest that adolescents and young adults account for half of all new STI cases in the U.S. every year [3].

To help identify programs effective in reducing these risks, since 2009, the U.S. Department of Health and Human Services has contracted with Mathematica Policy Research and its partner, Child Trends, to conduct an independent systematic review of the evidence base on programs to reduce teen pregnancy, STIs, and associated sexual risk behaviors. The review identifies, assesses, and rates the rigor of program impact studies and describes the strength of evidence supporting different program models. Findings are used to identify program models meeting the criteria for the HHS List of Evidence-Based Teen Pregnancy Prevention Programs.

Findings from the first review of the evidence, completed in spring 2010, were released in conjunction with the Office of Adolescent Health (OAH) Teen Pregnancy Prevention (TPP) program grant announcements. The findings were also highlighted in the 2010 State Personal Responsibility Education Program (PREP) grant announcement. In December 2010, Mathematica and Child Trends released a public call for studies to update the review with new research findings. Results from this update to the review are expected to be released in fall 2011.

OBJECTIVES

The objectives of the review are to:

1. Identify, assess, and rate the rigor of studies examining program impacts on teen pregnancy, STIs, and associated sexual risk behaviors.
2. Describe the strength of evidence supporting different teen pregnancy prevention program models.
3. Identify program models meeting the criteria for the HHS List of Evidence-Based Teen Pregnancy Prevention Programs.
4. Strengthen the evidence base by identifying key gaps in the literature and setting standards for study quality and evidence of program effectiveness.

Inclusion Criteria

TYPES OF PARTICIPANTS

The review considers studies on United States youth ages 19 or younger. Studies with a subsample outside of this age range are considered for review if the study establishes that the majority of sample members are 19 or younger. There is no lower bound on age.

TYPES OF INTERVENTIONS

Interventions may focus on a range of approaches to prevent teen pregnancy, such as encouraging teens to wait to have sex, providing information on contraception, teaching refusal skills, or discussing the health consequences of sexual activity. Studies of interventions lacking such a focus, including research on dropout prevention, job training, early childhood education, and home visiting for adolescent mothers are excluded from the review. Studies of state- or federal-policy changes, such as policies affecting access to contraception through Medicaid, are likewise excluded.

TYPES OF STUDIES

Studies must examine the effects of an intervention using quantitative data, statistical analysis, and hypothesis testing. They must also have been conducted or published since 1989.

TYPES OF OUTCOMES

Studies must measure program impacts on at least one measure of sexual risk behavior or its health consequences. Measures meeting this definition include those examining: sexual activity (initiation, frequency, number of partners); contraceptive use; sexually transmitted infections (STIs); pregnancies; or births. Most studies use self-reported measures, but biological measures of sexually transmitted infections and administrative data (for example, birth records) are also considered. Measures with limitations in terms of their quality or interpretation (for example, reports from males of their female partners' use of birth control pills or scales of behavioral risk and contraceptive use, which combine multiple measures into a single "black box" scale) are excluded from the review.

SEARCH STRATEGY

Studies are identified for review in four ways: reviewing published research syntheses, reviewing the websites of relevant research and policy organizations, issuing public calls for studies to solicit new and unpublished research, and conducting keyword searches of electronic databases.

1. REVIEW OF RESEARCH SYNTHESES

For the initial review of the evidence, the review team identified relevant studies by scanning the reference lists of 7 syntheses of research studies related to adolescent pregnancy prevention (see Table A.1 for list).

2. WEBSITES OF RELEVANT RESEARCH AND POLICY ORGANIZATIONS

Additional studies are identified by searching the websites of federal agencies and research or policy organizations with links to the topic of teen pregnancy prevention. For the initial review of the evidence, the review team searched the websites of nine such agencies or organizations (see Table A.2 for list).

3. CALL FOR STUDIES

New studies and unpublished studies of relevance are identified by the review team through periodic public calls for studies. For the first review of the evidence, a public call for studies was distributed in September 2009 to 43 research organizations, professional associations, and university-affiliated research centers (see Table A.3 for list). Authors were given approximately six weeks to submit materials.

4. KEYWORD SEARCH OF ELECTRONIC DATABASES

Additional studies are identified by conducting keyword searches of electronic citation databases. For the first review of the evidence, the review team coordinated with Mathematica's professional research librarians to conduct a search of 12 electronic databases (see Table A.4 for list) with the following keyword combination:

pregnancy OR pregnant OR "HIV" OR "AIDS" OR "STD" OR "sexually transmitted"

AND (prevention OR clinic) AND (adolescent OR teen*)*

AND (evaluation OR stud*) AND (effect* OR impact*)*

AND ("sex education" OR (sex AND education) OR abstinence)

ASSESSMENT OF METHODOLOGICAL QUALITY

Studies that meet the review screening criteria are each assessed for quality of research design and implementation. The assessments are conducted by a team of researchers from Mathematica and Child Trends, all of whom receive a full-day training on the evidence review and assessment protocol. Each individual impact study is assessed by two team members; the first member conducts a detailed review of the study following a protocol developed by Mathematica and approved by HHS; the second member checks and verifies the review for accuracy and completeness. Following the assessment, the team

members assign each impact study a quality rating of ***high***, ***moderate***, or ***low*** for the rigor and execution of its research design. The rating scheme was developed by Mathematica and approved by HHS prior to the first review of the evidence. In developing the scheme, Mathematica drew upon the evidence standards used by nine other evidence assessment projects or research and policy groups (see Table B.1 for list).

1. STUDY DESIGN

The highest study quality rating is reserved for randomized controlled trials (RCTs) and similar studies that randomly assigned subjects to the study's research groups. Studies using random assignment provide the strongest evidence that differences in the outcomes between the treatment and control groups can be attributed to the intervention. (Designs based on functionally random assignment, such as alternating based on last name, date of birth, or certain digits of an identification number, are also eligible for this highest rating.)

Quasi-experimental designs with an external comparison group are eligible for at best a moderate rating. In such studies, subjects are sorted into the research groups through a process other than random assignment; therefore, even if the treatment and comparison groups are well matched based on observed characteristics, they may still differ on unmeasured characteristics. We therefore cannot rule out the possibility that the findings are attributable to unmeasured group differences. The moderate study rating is also applied to random assignment designs that do not meet other criteria for the highest rating (that is, attrition or reassignment), as explained in more detail below.

Quasi-experimental designs *without* an external comparison group (for example, pre-post designs) are given a low study rating. These designs are not considered for either the high or moderate rating because they offer no credible means to assess what the sample's outcomes would have been absent the intervention—a necessary condition for obtaining an unbiased impact estimate. Quasi-experimental and random assignment studies that do not meet the other criteria for a high or moderate rating will also be assigned the lowest rating.

2. ATTRITION

In random assignment studies, a loss of study participants can bias the study's impact estimates by creating differences in the characteristics of the treatment and control groups. Bias can arise from overall attrition (the percentage of study participants lost among the total study sample) or differential attrition (the difference in attrition rates between the treatment and control groups).

Attrition is assessed against standards established by the U.S. Department of Education's What Works Clearinghouse (WWC). As seen in Figure B.1, the WWC standards recognize a trade-off between overall and differential attrition. Namely, for an expected level of bias, studies with a relatively low level of overall attrition will be able to meet standards with a relatively high level of differential attrition, whereas studies with a relatively high level of overall attrition will require a lower level of differential attrition to meet standards. Thus, the cut-off for an acceptable level of sample attrition is, appropriately, tied not only to the extent of overall attrition or differential attrition but rather to a combination of the two. For example, for studies with a relatively low overall attrition rate of 10 percent, the WWC

standard allows a rate of differential attrition up to approximately 6 percent. However, for studies with a higher overall attrition rate of 30 percent, the WWC standard requires a lower rate of differential attrition, at approximately 4 percent. Only random assignment studies meeting the standard for acceptable combinations of overall and differential attrition are considered for the highest study rating. Random assignment studies that do not meet these standards are considered for the moderate study rating.

For cluster randomized trials, in which individuals are assigned to treatment and control conditions in groups (for example, schools or classrooms), the review team first assesses the level of attrition for the clusters, or groups. If the combination of overall and differential attrition at the cluster level meets the attrition standards, the review team then assesses attrition at the sub-cluster (or individual) level. Random assignment studies with low attrition at the cluster level but high attrition at the sub-cluster level are assigned the moderate study rating. Cluster randomized trials will also receive a moderate rating if sample members were added during the intervention period—for example, if a study of a multiyear pregnancy prevention program for high school students added to the sample new students who transferred into the school the year after the program began.

The attrition standards are not applied to quasi-experimental studies, because these studies are evaluated on the basis of their final analytic samples (explained later), from which there is no attrition.

3. BASELINE EQUIVALENCE

In quasi-experimental comparison group studies and random assignment studies with high attrition, the use of well-matched treatment and comparison groups can minimize the bias in the impact estimates. Therefore, in order to receive the moderate study rating, quasi-experimental comparison group studies and random assignment studies with high attrition are required to demonstrate that the intervention and comparison groups were similar at baseline ($p > .05$, two-tailed test) on three key demographic characteristics: age or grade level, gender, and race/ethnicity. Studies are also required to establish baseline equivalence on at least one behavioral outcome measure (for example, rates of sexual initiation), unless the study sample was too young (that is, younger than age 14 or eighth grade) at baseline to expect that such behaviors were measured.

Only those outcomes for which baseline equivalence is established are considered for possible evidence of program effectiveness. For example, if a study examined program impacts on three relevant outcome measures—sexual initiation, contraceptive use, and pregnancy—but established baseline equivalence for only one of the three measures (sexual initiation), the study meets the criteria for a moderate study rating, but only the impact findings for that one outcome measure (sexual initiation) are considered for possible evidence of program effectiveness. Studies are also required to control for these measures in their analyses, to ensure that any marginal differences in outcome measures at baseline did not bias the impact estimates at follow up.

These baseline equivalence criteria are assessed on the study's final analysis sample. In some cases, studies assess equivalence for all youth who completed a baseline survey, but then present impact estimates for only a smaller subset of youth who completed a follow-up survey. These studies do not meet the baseline equivalence criteria of this review, because equivalence was not established for the smaller subset of youth on which the program impacts were based. Similarly, studies are not considered

for the moderate rating if they present baseline equivalence statistics separately for subgroups defined by age, gender, or race/ethnicity, without also establishing equivalence for the full analytic sample. Some studies, for example, present baseline equivalence statistics separately for males and females or for subgroups of older and younger youth, but not for the overall combined sample.

Random assignment studies that otherwise meet the criteria for the highest rating are not required to establish baseline equivalence, because randomization is expected to produce groups that are equivalent, on average, on both observed and unobserved characteristics. Nevertheless, randomization sometimes can produce chance differences between groups and, to meet the criteria for the highest study rating, random assignment studies that show evidence of statistically significant baseline differences on behavioral outcome measures or demographics (age, race/ethnicity, or gender) are required to control for these differences in their statistical impact analyses. Random assignment studies that do not control for statistically significant baseline differences are assigned the moderate rating.

4. REASSIGNMENT

In random assignment studies, deviation from the original random assignment (for example, moving youth from the treatment to the control group) can bias the study's impact estimates. Therefore, in order for a random assignment study to meet the criteria for the highest rating, the analysis has to have been performed on the sample as originally assigned. In order to receive a high rating, subjects cannot be reassigned, based on actual treatment they received, for reasons such as contamination, noncompliance, or level of exposure. Random assignment studies that somehow alter the original random assignment must establish baseline equivalence of their final analysis sample in order to be considered for a moderate study rating.

5. CONFOUNDING

In certain cases, a component of the research design or methods lines up exactly with the intervention being tested, undermining the credibility of attributing an observed effect to the intervention. For example, if a study assigns only one subject or group (for example, classroom or school) to the treatment or control condition, there is no way to distinguish the effects of the program from the particular effects of that one assigned subject or group. This can happen, for example, in quasi-experimental comparison group studies that estimate program impacts by comparing a single school or school district that implemented a pregnancy prevention program with a neighboring school or school district that did not have the program. In these cases, there is no way to distinguish the effects of the program from other characteristics of the particular school or district that implemented the program. A confounding factor can also arise from systematic differences in data collection methods for the treatment and comparison groups—for example, if program staff collects data from all subjects in the treatment group but an independent group of staff collects data from the control group. In this case, the mode of data collection cannot be separated from the effects of the intervention. Because the presence of such confounding factors severely weakens the credibility of a study's findings, a low rating is assigned to random assignment or quasi-experimental comparison group studies with either (1) only one subject or group in the treatment and control condition or (2) systematic differences in data collection procedures between the treatment and control groups.

DATA COLLECTION/EXTRACTION

All impact studies meeting the criteria for a high or moderate study quality rating are considered eligible for providing credible evidence of program impacts. For these eligible studies, the review team documents the impact estimate(s) for all relevant outcome measures, and this information is used to assess a program's evidence of effectiveness. Studies receiving a low rating are not subject to data collection and extraction, as the information provided in these studies is considered not to provide credible estimates of program impacts.

For each relevant impact estimate from an eligible impact study, the review team collects and records the following information: The name and description of the measure; the type of outcome the measure examined; the sample to which the impact estimate pertains (full sample or subgroup of interest defined by (1) gender or (2) sexual experience at baseline); the follow-up period to which it pertains; the point estimates of the intervention and comparison groups; the magnitude of the impact estimate; the reported statistical confidence interval or associated standard error of the estimate; the reported *p*-value or other associated test statistic; and whether the estimate is reported as statistically significant.

In the case of random assignment studies with multiple follow-up periods, this information is documented only for follow-up periods meeting the standard for low sample attrition. For follow-up periods not meeting the attrition standard, the information is treated as if it was based on a moderate quality study and documented only if the study establishes baseline equivalence for the analysis sample of that follow-up.

The review team documents all of this information as the author(s) reports it. For example, studies can report the magnitude of the impact estimates in many forms—as log-odds ratios, differences in probabilities, or effect size units—and the review team documents each magnitude as it is reported. For many studies, information on the impact estimates is not complete and the review team must document certain information as missing. Unfortunately, missing information is particularly common with respect to the magnitude and standard errors of the impact estimates, which makes it difficult to standardize the impact findings (for example, into effect size units) and compare them across different outcomes, different studies or programs, or against external benchmarks.

Based on the information collected and extracted from the eligible impact studies, the review team qualitatively describes the strength of evidence supporting each program model and identifies those programs meeting the criteria for the HHS List of Evidence-Based Teen Pregnancy Prevention Programs. To meet the HHS criteria, the program's supporting research study must show evidence of a positive, statistically significant impact on at least one priority outcome measure for either the full analytic sample or a subgroup defined by (1) gender or (2) sexual experience at baseline. The priority outcome measures are sexual activity (initiation; frequency; rates of vaginal, oral and/or anal sex; number of sexual partners), contraceptive use (consistency of use or one-time use, for either condoms or another contraceptive method), STIs, and pregnancy or birth. Statistical significance is assessed with a two-tailed hypothesis test and a specified alpha level of $p < .05$.

Although commonly featured in the literature, evidence from subgroups defined by sexual activity at follow-up are not considered when identifying programs for the HHS List of Evidence-Based Teen Pregnancy Prevention Programs. As with other endogenous subgroups that are defined by behavior emerging after the start of the program, the composition of those who are sexually active at follow up may be affected by program participation. As a result, even with an experimental design, the treatment and comparison groups within such subgroups may lack equivalence, leading to biased estimates of a program's impact for these groups.

MAIN FINDINGS

For the first review of the evidence, the search strategy identified approximately 1,000 potentially relevant studies for review. A total of 199 studies met the review screening criteria, and 93 studies received either a high or moderate study quality rating. From these 93 studies, 28 program models were identified as meeting the criteria for the HHS List of Evidence-Based Teen Pregnancy Prevention Programs. Of these 28 program models, 19 are supported by program impact studies that received high study quality ratings, and 9 are supported by studies that received moderate quality ratings (Tables C.1 and C.2). A majority of programs are supported by a single program impact study showing evidence of short- or long-term program impacts for the full study sample. The review team found no programs with evidence of sustained, full-sample impacts replicated across two or more high-quality studies.

CONFLICTS OF INTEREST

None.

REFERENCES

1. Centers for Disease Control and Prevention. "Youth Risk Behavior Surveillance—United States, 2009." *MMWR* 2010; 59(No. SS-5): 1–142.

2. Ventura, S.J. and B.E. Hamilton. "U.S. Teenage Birth Rate Continues to Decline." NCHS Data Brief, No. 58, February 2011.

3. Weinstock, H., Berman, S., and W. Cates, Jr. "Sexually Transmitted Diseases Among American Youth: Incidence and Prevalence Estimates, 2000." *Perspectives on Sexual and Reproductive Health*, vol. 36, no. 1, 2004, pp. 6-10.

Appendix A – Search Strategy

TABLE A.1: RESEARCH SYNTHESES

1. Advocates for Youth. (2008). *Science and Success,* 2nd edition. Washington, DC: Advocates for Youth.
2. Guide to Community Preventive Services. Prevention of HIV/AIDS, other STIs and pregnancy: Group-based abstinence education interventions for adolescents. (http://www.thecommunityguide.org/hiv/abstinence_ed.html).
3. Guide to Community Preventive Services. Prevention of HIV/AIDS, other STIs and pregnancy: Group-based comprehensive risk reduction interventions for adolescents. (http://www.thecommunityguide.org/hiv/riskreduction.html).
4. Kim, C. C., & Rector, R. (2008). "Abstinence education: Assessing the evidence." Washington, DC: The Heritage Foundation.
5. Kirby, D. (2007). *Emerging Answers 2007: Research Findings on Programs to Reduce Teen Pregnancy and Sexually Transmitted Diseases.* Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy.
6. Oringanje, C., Meremikwu, M. M., Eko, H., Esu, E., Meremikwu, A., & Ehiri, J. E. (2009). "Interventions for preventing unintended pregnancies among adolescents." *Cochrane Database of Systematic Reviews*, Issue 4.
7. Scher, L., Maynard, R. A., & Stagner, M. (2006). "Interventions intended to reduce pregnancy-related outcomes among adolescents." *Campbell Systematic Reviews*, Number 12.

TABLE A.2: RELEVANT WEBSITES

1. Advocates for Youth
2. Centers for Disease Control and Prevention (HIV/STD Prevention Research Synthesis)
3. Guttmacher Institute
4. Healthy Teen Network
5. National Abstinence Clearinghouse
6. National Abstinence Education Association
7. National Campaign to Prevent Teen and Unplanned Pregnancy
8. Sociometrics (Program Archive on Sexuality, Health, and Adolescence)
9. Child Trends (LINKS database)

TABLE A.3: CALL FOR STUDIES DISTRIBUTION LIST

Abt Associates
American Association for Marriage and Family Therapy
American Counseling Association
American Educational Research Association
American Evaluation Association
American Psychological Association, Society for Child and Family Policy and Practice Division
American Psychological Association, Society for Community Research and Action Division
American Public Health Association
Annie E. Casey Foundation
Association for Public Policy Analysis and Management
Association of Maternal and Child Health Programs
CDC Prevention Research Centers
CDC Division of Reproductive Health State and National Partners
Child Trends
CityMatch
ETR Associates
Ford Foundation
Guttmacher Institute
Healthy Teen Network
Heritage Foundation
National Association of County and City Health Officials
National Abstinence Clearinghouse
National Abstinence Education Association
National Association of Social Workers
National Campaign to Prevent Teen and Unplanned Pregnancy
National Council on Family Relations
National Institutes of Health Behavioral and Social Science Listserv
Office of Population Affairs
Oregon Social Learning Center
Pacific Institute for Research and Evaluation
Philliber Research Associates
Population Council
Prevention Research Center, Arizona State University
Prevention Research Center, Pennsylvania State University
RTI International
Sexuality Information and Education Council of the United States
Social Development Research Group, University of Washington
Society for Adolescent Medicine
Society for Prevention Research
Society for Research in Child Development
Society for Research on Adolescence
Sociometrics
WT Grant Foundation

TABLE A.4: KEYWORD SEARCH DATABASES

1. Academic Search Premier
2. CINAHL with Full Text
3. Cochrane Methodology Register
4. Cochrane Central Register of Controlled Trials
5. Cochrane Database of Systematic Reviews
6. Database of Abstracts of Reviews of Effect
7. Dissertation Abstracts
8. Education Research Complete
9. ERIC
10. MedLine
11. PsycInfo
12. SocINDEX with Full Text

Appendix B – Evidence Standards

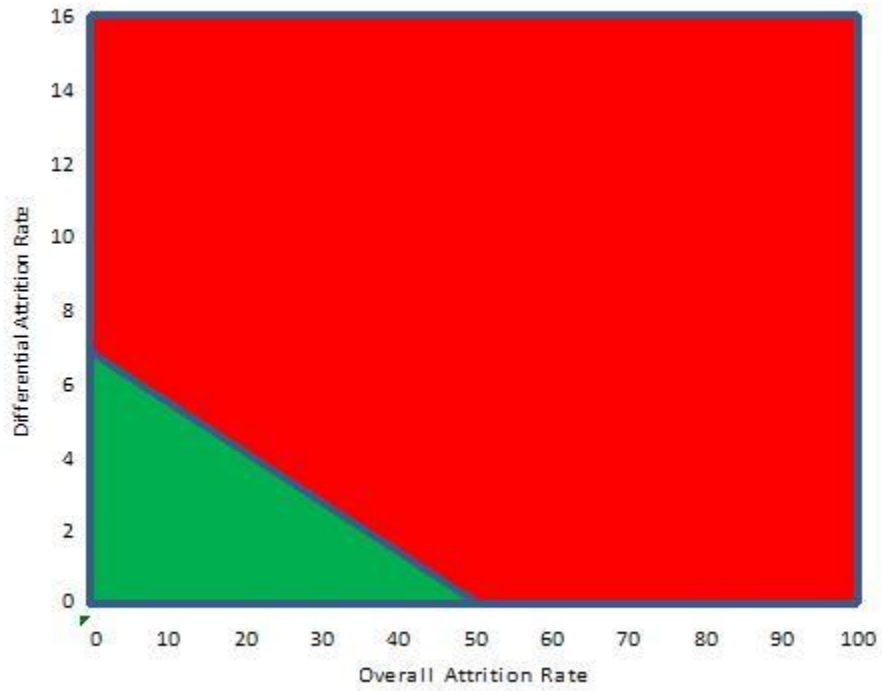
TABLE B.1: RELATED ORGANIZATIONS CONSULTED WHILE DEVELOPING REVIEW STANDARDS

1. Advocates for Youth. (2008). Science and Success, 2nd edition. Washington, DC: Advocates for Youth.
2. Blueprints for Violence Prevention
3. Centers for Disease Control and Prevention, HIV/AIDS Prevention Research Synthesis
4. Child Trends (LINKS Database)
5. Kirby, D. (2007). Emerging Answers 2007: Research Findings on Programs to Reduce Teen Pregnancy and Sexually Transmitted Diseases. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy.
6. National Registry of Evidence-Based Programs and Practices
7. Scher, L., Maynard, R. A., & Stagner, M. (2006). "Interventions intended to reduce pregnancy-related outcomes among adolescents." Campbell Systematic Reviews, Number 12.
8. Sociometrics (Program Archive on Sexuality, Health, and Adolescence)
9. What Works Clearinghouse

TABLE B.2 – SUMMARY OF STUDY QUALITY RATINGS

Criteria Category	High Study Rating	Moderate Study Rating	Low Study Rating
1. Study design	Random or functionally random assignment	Quasi-experimental design with a comparison group; random assignment design with high attrition or reassignment	Does not meet criteria for high or moderate rating
2. Attrition	What Works Clearinghouse standards for overall and differential attrition	No requirement	Does not meet criteria for high or moderate rating
3. Baseline equivalence	Must control for statistically significant baseline differences	Must establish baseline equivalence of research groups and control for baseline outcome measures	Does not meet criteria for high or moderate rating
4. Reassignment	Analysis must be based on original assignment to research groups	No requirement	Does not meet criteria for high or moderate rating
5. Confounding factors	Must have at least two subjects or groups in each research group and no systematic differences in data collection methods	Must have at least two subjects or groups in each research group and no systematic differences in data collection methods	Does not meet criteria for high or moderate rating

FIGURE B.1 – STANDARD FOR SAMPLE ATTRITION IN ASSIGNING STUDY QUALITY RATING



Source: What Works Clearinghouse. (2008). Procedures and Standards Handbook, Version 2. Washington, DC: U.S. Department of Education.

Appendix C – Evidence of Effectiveness

TABLE C.1 – SUMMARY OF EVIDENCE

	High-Quality Impact Study with Replicated Impact	High-Quality Impact Study with Sustained Impact	High-Quality Impact Study with Short-Term Impact	High-Quality Impact Study with Subgroup Impact
Study Quality	High	High	High	High
Sample with Positive Impacts	Full sample	Full sample	Full sample	Subgroup
Duration of Impacts	Year or more	Year or more	Less than year	Any
Replicated	Yes	No	Yes or no	Yes or no
Number of Programs	0	6	10	3

	Moderate-Quality Impact Study with Replicated Impact	Moderate-Quality Impact Study with Sustained Impact	Moderate-Quality Impact Study with Short-Term Impact	Moderate-Quality Impact Study with Subgroup Impact
Study Quality	Moderate	Moderate	Moderate	Moderate
Sample with Positive Impacts	Full sample	Full sample	Full sample	Subgroup
Duration of Impacts	Year or more	Year or more	Less than year	Any
Replicated	Yes	No	Yes or no	Yes or no
Number of Programs	0	2	3	4

TABLE C.2 – LIST OF HHS EVIDENCE-BASED PROGRAMS BY EVIDENCE CATEGORY

Program	Evidence Category
Aban Aya Youth Project	Moderate-quality study with subgroup impact
Adult Identity Mentoring (Project AIM)	High-quality study with short-term impact
All4You!	Moderate-quality study with short-term impact
Assisting in Rehabilitating Kids (ARK)	High-quality study with sustained impact
Be Proud! Be Responsible!	High-quality study with short-term impact
Be Proud! Be Responsible! Be Protective!	High-quality study with short-term impact
Becoming a Responsible Teen (BART)	High-quality study with sustained impact
Children's Aid Society (CAS)—Carrera Program	High-quality study with sustained impact
¡Cuídate!	High-quality study with short-term impact
Draw the Line/Respect the Line	High-quality study with subgroup impact
FOCUS	High-quality study with subgroup impact
Horizons	High-quality study with short-term impact
It's Your Game: Keep it Real	Moderate-quality study with subgroup impact
Making a Difference!	High-quality study with short-term impact
Making Proud Choices!	High-quality study with short-term impact
Project TALC	High-quality study with sustained impact
Promoting Health Among Teens! Abstinence Only Intervention	High-quality study with sustained impact
Promoting Health Among Teens! Comprehensive Abstinence and Safer Sex Intervention	High-quality study with short-term impact
Reducing the Risk	Moderate-quality study with subgroup impact
Rikers Health Advocacy Program (RHAP)	Moderate-quality study with short-term impact
Safer Sex	Moderate-quality study with short-term impact
Raising Healthy Children	Moderate-quality study with sustained impact
Sexual Health and Adolescent Risk Prevention (SHARP)	Moderate-quality study with sustained impact
SiHLE	High-quality study with short-term impact
Sisters Saving Sisters	High-quality study with sustained impact
Teen Health Project	Moderate-quality study with subgroup impact
Teen Outreach Program	High-quality study with subgroup impact
What Could You Do?	High-quality study with short-term impact