

# HHS List of Evidence-Based Program Models that Impact Rates of Teen Pregnancy, Sexually Transmitted Infections, or Other Sexual Risk Behaviors:

## Recent Review Process and Findings Frequently Asked Questions

April 20<sup>th</sup>, 2012

The U.S. Department of Health and Human Services (HHS) issued a contract to conduct a systematic review of the evidence base for programs that have shown evidence of preventing teen pregnancy, reducing the rates of sexually transmitted infections, and impacting rates of associated sexual risk behaviors. The review process was approved by a workgroup comprised of representatives from HHS, including the Administration for Children and Families (ACF), the Centers for Disease Control and Prevention (CDC), the Office of Adolescent Health (OAH), and the Office of the Assistant Secretary for Planning and Evaluation (ASPE). Mathematica Policy Research, Inc. conducted this review process with their subcontractor, Child Trends, and has a contract to maintain and update the evidence base as needed.

The results of this review have been used to produce the *HHS List of Evidence-Based Teen Pregnancy Prevention Program Models*<sup>1</sup>, which is a listing of program models with evidence of effectiveness in reducing teen pregnancy, sexually transmitted infections, or associated sexual risk behaviors. These programs reflect a range of approaches that exist in the field (e.g., youth development, comprehensive sex education, abstinence, HIV/STI education, etc.).

The following are a compilation of frequently asked questions (FAQs) that address information about the established review procedures and criteria, updates from the latest round of reviews, and a summary of findings to date. More detailed information on the specific process and criteria used to conduct the review can be found at (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/eb-programs-review-v2.pdf>)

### GENERAL QUESTIONS ABOUT THE REVIEW

- 1. The current HHS criteria for evidence of program effectiveness have no requirement for evidence of sustained impact (for example, impacts on short-term contraceptive use versus longer-term impacts on pregnancy deterrence). Will this be reconsidered in future updates of the review?**

Yes, HHS continues to consider many factors influencing the strength of the review criteria. As stated in the Review Protocol 2.0 (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/eb-programs-review-v2.pdf>), numerous dimensions were considered

---

<sup>1</sup> The HHS List of Evidence-Based Teen Pregnancy Prevention Program Models is a listing of programs with evidence of effectiveness for impacting rates of pregnancy, STIs, or sexual risk behaviors and be found at <http://www.hhs.gov/ash/oah/oah-initiatives/tpp/programs.html>.

when determining the criteria for program effectiveness during the initial round of reviews. These included:

- Quality of the supporting study or studies' design – e.g., did the study use a random assignment design, or a strong quasi-experimental design?
- Statistical significance and direction (i.e. were the findings positive or negative?) of each supporting study's impact estimate(s)
- Type of outcome measure(s) with statistically significant impact estimate(s)
- Number and proportion of the statistically significant impact estimate(s)
- Duration of the statistically significant impact estimate(s)
- Sample(s) with statistically significant impact estimate(s)—full sample or specified subgroup(s)
- Whether statistically significant impact estimate(s) had been replicated across multiple studies

From this list of possible dimensions, HHS specified the dimensions that would be used to categorize a program model as "evidence-based." As the research evidence on teen pregnancy prevention programs expands, HHS, with the help of Mathematica and expert consultants, will be determining areas that should be updated.

## **2. Did the studies have to appear in peer-reviewed journals in order to be included in the review?**

No. The review was not limited to peer-reviewed journal articles. The review also included studies reported as part of book chapters, government reports, unpublished manuscripts, or other documents.

By not restricting the pool of eligible studies to peer review publications, we are able to identify more recently evaluated studies (the publication process can be quite lengthy). In addition, not all peer-reviewed publication venues are the same in terms of quality of the review. Instead, we focused on the quality of the evaluation study and assessed the impacts based upon the established review criteria. We required authors of unpublished reports to provide complete information needed to assess the quality of the evaluation study and its outcomes.

## **3. Why were programs that produced impacts at 6 months but not 12 months deemed as effective? Why were programs with only 3 months impact deemed effective?**

The HHS criteria for evidence effectiveness required studies to demonstrate that program participants experienced a period of reduced risk for teen pregnancy, sexually transmitted infections, or associated sexual risk behaviors. A reduction in sexual risk behaviors, even over only a relatively short-term period of 3 or 6 months, met this standard and may have longer-term benefits for adolescent health and well-being.

#### **4. How many program models appear on the current HHS List of Teen Pregnancy Prevention Program Models?**

During the first review, the *HHS List of Evidence-Based Teen Pregnancy Prevention Programs* included 28 program models. The program models represent a range of different program approaches, including abstinence, comprehensive sex education, HIV/STI prevention, and youth development approaches. This list is found on the HHS website - (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/tpp-database.html>).

The first update of the review adds three new programs models to the list, for a total of 31 models. The three additional program models that have been identified are:

- o *Heritage Keepers*: an abstinence curricula evaluated with 7<sup>th</sup>-9<sup>th</sup> grade students.
- o *Respeto/Proteger*: an HIV Prevention Program evaluated with Latino teen parents.
- o *Safer Choices*: a school-based, comprehensive sex education model evaluated with 9<sup>th</sup> grade students.

#### **5. Were there any changes to the criteria used to conduct the latest round of reviews? Will there be changes in the future?**

No, the latest round of reviews did not incorporate any changes to the review criteria. However, it is possible that the criteria may be strengthened over time in order to align the criteria with advances in the field and with other established systematic review processes. HHS will continue to consult with leading experts in the fields of evaluation design, research methodology, and systematic reviews to help enhance the review process in the future.

#### **6. What years did the most recent round of reviews cover? How many studies were identified?**

The most recent round of reviews included studies published from January 2009 through December 2010. Additionally, a call for new studies was also made in December 2010 and closed in February 2011. During the update to the review, 90 new relevant studies were identified through the search strategy. Of these, a total of 26 studies met the review screening criteria, 10 studies received either a high or moderate quality study rating<sup>2</sup>, and 3 studies supported program models that were added to the list because they demonstrated impacts on key behavioral outcomes.

#### **7. How has the search strategy changed?**

The search strategy originally identified new studies 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 strategies. Under the more recent round of reviews, the search strategy has included scanning relevant research journals, and reviewing professional conference proceedings.

---

<sup>2</sup> For more information on how studies were assigned high or moderate-quality ratings, please reference Question 10 or the website.

## 8. The intervention implementation reports have changed. What information is new?

We have expanded upon the level of detailed provided in the *evaluation* sections of the intervention implementation reports. These changes include:

- **Enhanced discussion of the study rating.** Each report clearly indicates whether program model received a high or a moderate rating based on the supporting studies.
- **Detailed explanation of the study design.** In order to help advance the field and provide information about the initial evaluation, each report provides a detailed discussion of the study design.
- **Thorough explanation of the findings that met the criteria for evidence for the review.** The findings behind supporting studies have been expanded to include an explanation of the findings so that multiple audiences could determine the impacts found for each program, as well as areas where there were no impacts.
- **An explanation of non-sexual behavior outcomes, where applicable.** The updated reports indicate whether non-sexual behavior outcomes were measured in each study. However, it is important to note that these outcomes were *not* considered for the review. The purpose of this explanation is to provide information on other outcomes that these programs try to address.

## REVIEW PROCEDURES AND CRITERIA

### 9. What criteria were used to conduct the review?

In developing the review criteria, HHS drew upon evidence standards used by several well established evidence assessment projects and research and policy groups, such as the *What Works Clearinghouse*, *Blueprints for Violence Prevention*, and *the National Registry of Evidence-Based Programs and Practices*. Based on standards utilized in these other processes, this review defined the criteria for the quality of an evaluation study and the strength of evidence for a particular intervention. Using these criteria, the Department then defined a set of rigorous standards an evaluation must meet in order for a program to be considered effective and therefore eligible for funding as an evidence-based program.

Detailed information on the technical approach and criteria used to conduct the review is available on the OAH website (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/eb-programs-review-v2.pdf>). The criteria utilized for the first update to the review are the same as those applied during the initial round of the review process.

### 10. How were high and moderate quality study ratings defined?

The high study quality rating was reserved for randomized controlled trials with low rates of sample attrition, no reassignment of sample members, no systematic differences in data collection between the research groups, and at least one subject or group (school, classrooms, etc.) in both the treatment and control conditions. The moderate study quality rating was considered for studies using quasi-experimental designs and for randomized controlled trials that did not meet all the review criteria for a high quality rating. To meet the criteria for a moderate study quality rating, a study had to demonstrate equivalence of the intervention and comparison groups on race, age,

and gender; report no systematic differences in data collection between the research groups; and have at least one subject or group (school, classroom, etc.) in both the intervention and comparison conditions. Studies based on samples of youth ages 14 or older also had to demonstrate equivalence of the intervention and comparison groups on at least one behavioral outcome measure. More detailed information on the review criteria is available on the OAH website (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/eb-programs-review-v2.pdf>).

**11. Who was part of the review panel?**

The initial review panel consisted of a team of 14 researchers from Mathematica, Child Trends, and Concentric Research and Evaluation, all of whom participated in a full-day training on the evidence review and protocol for assessing each impact study. Each individual impact study was assessed by two team members; the first member conducted a detailed review of the study following a protocol developed by Mathematica and approved by the HHS interagency work group; the second member assessed and verified the review for accuracy and completeness.

**12. Were the outcomes of interest the actual behavior of teens or teens' intentions?**

The review was limited to measures of sexual risk behavior and its health consequences. Measures meeting this definition include those examining: sexual activity (e.g., initiation, frequency, or number of partners), contraceptive use, sexually transmitted infections, pregnancies, or births. The review did not consider measures of teens' attitudes or intentions.

**13. Did you only review programs if they have had a randomized controlled trial evaluation?**

No. In addition to studies that used a randomized design, the review also considered quasi-experimental studies that did not employ random assignment.

**14. Did you look only at U.S. studies?**

Yes. The review was limited to studies of programs serving youth in the United States.

**15. Did the age criterion of 19 or younger refer to the age of participants at the time of initial intervention or the maximum age of program participants? In other words, were studies reviewed in which a program participant may have turned 20 during the study period?**

The age criterion of 19 or younger refers to age at the start of the intervention. Participants may have been older than 19 during the study period or when outcome measures were assessed.

**16. Did you review research on clinic-based programs?**

Yes, research on clinic-based programs was included in the review. This includes research on individualized programs that were integrated directly into existing clinic services, as well as research on small group training or counseling sessions offered as separate add-on services in a clinic setting. In order to be eligible, clinic-based programs had to emphasize both education and clinical services.

**17. In studies that reported multiple outcome measures, did you (1) prioritize the "best" outcome and focus only on that outcome or (2) collect information for all outcomes and meta-analyze across them?**

For studies examining multiple outcome measures, evidence of a statistically significant positive program impact on any one outcome was considered sufficient for meeting the HHS criteria for evidence of effectiveness. The review team did not rank order or prioritize the measures or average impacts across them.

**18. How did you handle outcome measures of poor or questionable quality?**

Measures with serious limitations in terms of their validity or interpretation were excluded. For example, the review did not consider reports from males of their female partners' use of birth control pills, or scales of behavioral risk that combine multiple measures into a single outcome.

**19. Were any subgroups considered for the review?**

Yes. In addition to findings for the full study sample, the review also considered findings for two subgroups: (1) gender and (2) sexual experience at baseline. The review considered the same outcome measures for these subgroups as for the full study sample—namely, measures of sexual risk behavior and its health consequences.

**20. Why wasn't race/ethnicity included as one of the priority subgroups?**

The subgroup assessment was limited to address concerns about "multiple comparisons" or "multiple hypothesis testing." The issue is that as you increase the number of subgroups examined in a particular study, you also increase the probability of finding a statistically significant impact just by chance. To address this issue, we chose to limit the number of subgroups considered for providing evidence of effectiveness. In selecting these subgroups, there were many relevant options to consider, such as race/ethnicity, gender, sexual experience, socioeconomic status, family structure, and many others. HHS ultimately chose to focus on gender and baseline sexual experience as the two subgroups to be considered by the review. Moreover, in many studies, the sample sizes were too small to assess impacts separately by race/ethnicity.

## REVIEW FINDINGS

### **21. Is there a list of studies that were reviewed and did not meet the HHS criteria for evidence of effectiveness?**

A full list of the studies reviewed is available in a PDF document contained at: [www.hhs.gov/ash/oah/oah-initiatives/tpp/PPRER\\_Studies\\_Table\\_formatted\\_Final.pdf](http://www.hhs.gov/ash/oah/oah-initiatives/tpp/PPRER_Studies_Table_formatted_Final.pdf).

### **22. Where can I find information about the effect sizes for each of the relevant behavioral outcomes?**

For each of the 31 program models that met the HHS criteria for evidence of effectiveness, the review team created a brief intervention implementation report describing key features of the program and the supporting research evidence, including information on the outcome measures affected. These reports are available on the OAH website (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/programs.html>). The review team did not report the magnitude of the impacts across programs or rank programs by the number of outcomes affected. This is because some studies did not report effect sizes or provide sufficient information to be able to calculate an effect size. We are in the process of identifying and calculating effect sizes for future rounds of the review and will include this information on intervention implementation reports in the future.

### **23. Why does the HHS list of evidence-based programs differ from similar lists I've seen from other groups?**

Each evidence review uses a slightly different set of procedures and criteria. Although there is usually overlap across lists, differences in the criteria used to screen and assess studies may lead to some difference in the list of programs identified as evidence-based.

As previously mentioned, in conducting the review, HHS drew upon evidence standards used by several well established evidence assessment projects and research and policy groups, such as the *What Works Clearinghouse*, *Blueprints for Violence Prevention*, and the *National Registry of Evidence-Based Programs and Practices*.

### **24. Why aren't there more programs on the list for high-risk populations such as Latinos and Native Americans?**

We did not start out by looking for programs targeted to any specific population. Rather, the goal of the review was to identify programs with the strongest evidence of effectiveness. If certain high-risk populations are underrepresented on the list, it points to a gap in the underlying research literature and need for additional research to identify effective programs for these groups. On the OAH website, the intervention implementation reports indicate which populations were included in the sample for each program evaluation.

### **25. Are there any abstinence programs on the list?**

The 31 program models represent a range of different program approaches, including abstinence, comprehensive sex education, HIV/STI prevention, and youth development approaches. You can also find more information on the programs on the OAH website.

There is a section of the website entitled "Evidence Based Programs" (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/programs.html>). This section of the website lists the 31 programs currently on the HHS List of Evidence-Based programs and provides links to more detailed information about each program.

**26. What were the racial/ethnic samples of the studies reviewed and the 31 programs that met the HHS criteria for evidence of effectiveness?**

Among the 31 programs meeting the HHS criteria for evidence of effectiveness, 16 were evaluated with predominately African American research samples, 8 were evaluated with mixed-race samples with no majority group, 4 were evaluated with predominately Latino research samples, and 3 were evaluated with research samples of predominately white adolescents. To the extent that certain population subgroups are under-represented on this list, it reflects the current state of the research literature and need for additional research to identify effective programs for under-represented groups. The review team did not collect comparable information for the broader list of all 199 studies reviewed.

**27. I don't understand or I disagree with the rating a study received.**

There is a section on the OAH website entitled "Review Protocol 2.0" (<http://www.hhs.gov/ash/oah/oah-initiatives/tpp/eb-programs-review-v2.pdf>) which provides a detailed description of the process used to determine the study ratings and an explanation of the criteria. This information should help explain the rating given to any particular study.

**28. I'm looking for a particular program on the HHS list of evidence-based programs, but it does not show up. Why was this particular program left off the list?**

There are four reasons a program may not appear on the list. First, a program may not be included because it has not been evaluated. Second, a program may not be included because even though it was evaluated, it did not meet screening criteria. Third, a program may not be included because its evaluation was not sufficiently rigorous. Fourth, a program may not be included because it did not provide evidence of positive impacts on one of the key outcome measures for either the full study sample or a priority subgroup.

## **FUTURE PLANS FOR UPDATING THE REVIEW**

**29. How often will the review be updated?**

The review will be updated on a periodic basis in order to both identify new evaluation research and to improve upon the review criteria as best practice and standards for evaluation methodology and conducting systematic reviews evolve.

**30. When new programs are deemed evidence-based, will they be considered eligible for funding by OAH and ACF under future replication grant funding opportunities if new money becomes available?**

It is up to the individual program offices to determine whether grantees will be able to replicate programs that are added to the list in the future. Though the evidence-based list will be used to inform any future funding announcements, program offices may consider additional factors beyond whether a program meets criteria for being evidence-based when deciding which programs are eligible for replication. For example, based on experience with the first round of teen pregnancy prevention initiative grant funding, program offices may focus on replicating those evidence-based program models that are implementation-ready and have training and support materials available.

**31. Will subsequent rounds of the evidence review reassess research from new studies of program models that have already been reviewed and recognized as evidence-based?**

Yes, each round of the evidence review will be open to any new studies or analyses of program models that have already been reviewed and recognized as evidence-based. Examples include replication studies conducted with different populations or in different settings, as well as new extensions or re-analyses of existing studies and data. However, Mathematica will not review research that was previously reviewed during the first review of the evidence base.

**32. In the future, if the evidence criteria are changed to become more rigorous, will existing programs on the list for replication that do not meet the new higher standard of evidence be removed from the list?**

Yes, if the evidence criteria are changed to become more rigorous, it is possible that program models that do not meet the more rigorous standard of evidence will be removed from the list.

**33. Are there plans to review any program models that have not been evaluated?**

No, the review is limited to program models that have been subject to a quantitative program impact evaluation. The review will assess any new quantitative impact evaluations as they become available. These include evaluations of program models that have been previously evaluated, studies of programs that have been adapted, or of new program models that had not previously been tested. As part of the continuing process, Mathematica will be reviewing new studies on a periodic basis.

**34. Will the next round of the evidence review consider adolescent health-risk behaviors beyond those related to pregnancy and sexual activity, such as teen dating violence or, self-harm?**

No, HHS has determined that program models eligible for replication must demonstrate evidence of a positive, statistically significant impact on at least one of the following outcomes:

- 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
- Pregnancy or birth

It is possible that programs effective in influencing these behaviors also affect other types of adolescent health-risk behaviors. However, to be included in the review, programs must examine program impacts on at least one measure of sexual risk behavior or its health consequences.

**35. Are there any plans to review programs that have been implemented with young adults over the age of 19?**

Yes, some studies included in the review contain sample participants above age 19. These studies are eligible for review as long as the majority of the participants are 19 and under.

Currently, there are no plans to review programs that have been implemented with youth exclusively above age 19, as the focus of the current review is on youth up to age 19. However, the age criterion of 19 or younger refers to the age of the youth at the start of the intervention. Participants may older than 19 during the study period or when outcomes are assessed.

**36. Will there be future OAH or ACF grant announcements or calls for proposals?**

It is possible that there may be future OAH or ACF grant announcements for additional grantees, depending on the availability of funds that are appropriated for these programs in the future.

**37. Can I submit a study for review?**

Yes, you may submit a study to [PPRER@mathematica-mpr.com](mailto:PPRER@mathematica-mpr.com). However, new studies will only be considered during the next round of reviews.

*Several of these questions were discussed during a webinar held in December 2010 to provide more detailed guidance on the review to the general public. The audio recording, transcript, and slides from the webinar can be accessed at (<http://www.hhs.gov/ash/oah/oah-initiatives/webinars/evreview2011audio.wmv>).*