

# Adolescent Pregnancy Prevention: Meta-Analysis of Federally Funded Program Evaluations

Randall Juras, PhD, Emily Tanner-Smith, PhD, Meredith Kelsey, PhD, Mark Lipsey, PhD, and Jean Layzer, MEd

**Background.** Beginning in 2010, the US Department of Health and Human Services (HHS) funded more than 40 evaluations of adolescent pregnancy prevention interventions. The government's emphasis on rigor and transparency, along with a requirement that grantees collect standardized behavioral outcomes, ensured that findings could be meaningfully compared across evaluations.

**Objectives.** We used random and mixed-effects meta-analysis to analyze the findings generated by these evaluations to learn whether program elements, program implementation features, and participant demographics were associated with effects on adolescent sexual risk behavior.

**Search Methods.** We screened all 43 independent evaluation reports, some of which included multiple studies, funded by HHS and completed before October 1, 2016. HHS released, and our team considered, all such studies regardless of favorability or statistical significance.

**Selection Criteria.** Of these studies, we included those that used a randomized or high-quality quasi-experimental research design. We excluded studies that did not use statistical matching or provide pretest equivalence data on a measure of sexual behavior or a close proxy. We also excluded studies that compared 2 pregnancy prevention interventions without a control group. A total of 44 studies from 39 reports, comprising 51 150 youths, met the inclusion criteria.

**Data Collection and Analysis.** Two researchers extracted data from each study by using standard systematic reviewing and meta-analysis procedures. In addition, study authors provided individual participant data for a subset of 34 studies. We used mixed-effects meta-regressions with aggregate data to examine whether program or participant characteristics were associated with program effects on adolescent sexual risk behaviors and consequences. To examine whether individual-level participant characteristics such as age, gender, and race/ethnicity were associated with program effects, we used a 1-stage meta-regression approach combining participant-level data (48 635 youths) with aggregate

data from the 10 studies for which participant-level data were not available.

**Main Results.** Across all 44 studies, we found small but statistically insignificant mean effects favoring the programs and little variability around those means. Only 2 program characteristics showed statistically reliable relationships with program effects. First, gender-specific (girl-only) programs yielded a statistically significant average effect size ( $P < .05$ ). Second, programs with individualized service delivery were more effective than programs delivering services to youths in small groups ( $P < .05$ ). We found no other statistically significant associations between program effects and program or participant characteristics, or evaluation methods. Nor was there a statistically significant difference in the mean effect sizes for programs with previous evidence of effectiveness and previously untested programs.

**Conclusions.** Although several individual studies reported positive impacts, the average effects were small and there was minimal variation in effect sizes across studies on all of the outcomes assessed. Thus, we were unable to confidently identify which individual program characteristics were associated with effects. However, these studies examined relatively short-term effects and it is an open question whether some programs, perhaps with distinctive characteristics, will show longer-term effects as more of the adolescent participants become sexually active.

**Public Health Implications.** The success of a small number of individualized interventions designed specifically for girls in changing behavioral outcomes suggests the need to reexamine the assumptions that underlie coed group approaches. However, given the almost total absence of similar programs targeting male adolescents, it is likely to be some time before evidence to support or reject such an approach for boys is available. (*Am J Public Health*. 2019;109:e1–e8. doi:10.2105/AJPH.2018.304925)

## PLAIN-LANGUAGE SUMMARY

We used study-level data from 44 federally funded impact evaluations of adolescent pregnancy prevention programs conducted between 2010 and 2016 and individual participant data from 34 of those studies to investigate whether the effects of these programs on sexual behavior and pregnancy are associated with isolated program or participant characteristics. Evidence on what works for whom would help program

developers design more effective programs and help practitioners select and implement programs most appropriate to the characteristics of their local youth populations. However, we found only small average program effects with little variability across programs, yielding limited evidence that any isolated program or participant characteristic was associated with larger effects on the pregnancy and sexual risk behavior outcomes

examined. The exceptions were some indication that girls-only single-gender programs and programs utilizing individualized service delivery might be effective. We recommend further research with longer follow-up periods to investigate whether these prevention programs show more positive effects as the adolescent participants mature and experience increased risk as they become more sexually active.

In 2010, the US Department of Health and Human Services (HHS) funded more than a hundred grants for programs to prevent adolescent pregnancy through a tiered funding strategy. A major portion of the funding went to programs that had demonstrated evidence of effectiveness according to a systematic review sponsored by HHS (tier 1).<sup>1</sup> The remainder went to previously untested programs (tier 2). HHS required a subset of the tier 1 grantees (on the basis of the size of the grant) and all tier 2 grantees to participate in rigorous independent evaluations, resulting in more than 40 new program evaluations across both tiers.<sup>2</sup>

These evaluations provided a unique opportunity to generate evidence on what works and for whom because of their rigor (grantees received intensive technical assistance by a federal evaluation contractor), consistently reported outcomes (a common core of survey items and behavioral outcomes was used in all of the evaluations), transparency (the agency agreed to release evaluation findings regardless of their results), and quality and fidelity of implementation (92% of all sessions observed by an independent facilitator were rated as very high or high quality, and 95% were implemented with high fidelity to the specified program model).

We applied meta-analytic techniques to systematically summarize the findings generated by these evaluations with a particular focus on identification of any program elements, implementation features, and participant demographics that were associated with program effects on adolescent sexual risk behavior and consequences.

## METHODS

This report adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement on standardized reporting for meta-analyses.<sup>3</sup> The methods were based on a protocol developed before data collection and analysis.<sup>4</sup> A small number of deviations from that protocol are described in Appendix A.

We screened all 43 reports of independent evaluations funded by HHS and completed before October 1, 2016. All such reports are available on HHS's Web site and at the National Library of Medicine.<sup>5</sup> Findings from

a large subset of the studies were published in the *Journal*.<sup>2</sup>

Two doctoral-level researchers (E. T. -S. and another researcher) independently conducted eligibility screening. To be eligible for inclusion, studies had to meet each of the following criteria:

1. evaluated an adolescent pregnancy prevention program,
2. included a comparison condition (i.e., no treatment or an alternate treatment not related to adolescent pregnancy),
3. used a randomized design or a controlled quasi-experimental design that matched participants or provided baseline equivalence data on a measure of previous sexual behavior,
4. assigned at least 10 participants to each condition, and
5. measured and reported on at least 1 sexual risk behavior or closely related outcome.

## Aggregate Data Extraction

We extracted aggregate data from each study by using a standardized coding protocol.<sup>6</sup> Master's- or doctoral-level researchers double-coded data from each report after several weeks of training followed by weekly coding meetings.

We extracted data on behavioral outcomes, program elements, implementation features, participant characteristics, and study methods from each report. We extracted attendance and retention data for each program from a database maintained by the Office of Adolescent Health. These variables are briefly described here, with detailed definitions provided in Appendix A. The full data extraction form is provided as Appendix B. The resulting data set is not publicly available at this time.

**Behavioral outcomes.** We coded sexual behavior outcomes in 9 categories:

1. ever had sex,
2. recent sexual activity,
3. recent unprotected sexual activity (i.e., without condoms or birth control),

4. number of sexual partners,
5. number of sexual experiences,
6. proportion of sexual experiences that were unprotected,
7. sexually transmitted infections,
8. ever pregnant or parent, and
9. recent pregnancy or parenting.

Most studies reported binary measures for these outcomes, so we used the log odds ratio (LOR) as the primary effect size metric with values greater than 0 indicating beneficial program effects relative to the comparison condition (e.g., lower odds of sexual behavior). For outcomes measured on a continuous scale (e.g., mean number of sexual partners), we used the small-sample corrected standardized mean difference effect size metric (Hedges' *g*).<sup>7</sup>

For studies ( $k = 20$ ) that used cluster assignment designs, we multiplied the standard errors of effect size estimates by the square root of the design effect to account for the nesting of participants within clusters (e.g., schools).<sup>8</sup>

**Study-level moderators.** We extracted a large set of moderators related to program elements and implementation features (itemized in Table 1): the primary program focus (e.g., youth development), program components (e.g., role plays), group size (e.g., small groups), group composition (e.g., mixed gender), gender-specific targeting (e.g., programs designed for girls only), program contact (e.g., daily), setting (e.g., classrooms), delivery personnel (e.g., classroom teachers), fidelity to the program design (i.e., based on implementation monitoring data), and participant engagement (e.g., attendance rate).

We also extracted 4 types of participant demographic characteristics at the study level: gender composition (percentage boys), racial/ethnic composition (percentage Black, White, Hispanic), average age, and risk level (based on rates of recent sexual activity).

## ABOUT THE AUTHORS

Randall Juras is with Abt Associates, Durham, NC. Emily Tanner-Smith and Mark Lipsey are with Peabody Research Institute, Vanderbilt University, Nashville, TN. Meredith Kelsey is with Abt Associates, Cambridge, MA. Jean Layzer is with Belmont Research Associates, Belmont, MA.

Correspondence should be sent to Randall Juras, 5001 S Miami Blvd #210, Durham, NC 27510 (e-mail: randall\_juras@abtassoc.com). Reprints can be ordered at <http://www.ajph.org> by clicking the "Reprints" link.

This article was accepted December 10, 2018.

doi: 10.2105/AJPH.2018.304925

*Study methods and risk of bias.* Finally, to assess for potential risk of bias in effect size estimates, we examined the empirical influence of 3 study-level moderators related to aspects of the study methods especially capable of creating bias: assignment to conditions (randomized vs matched), overall attrition rate, and differential attrition rates for the treatment and control conditions. Another moderator that might influence effect sizes was also included: an active control or comparison group (an alternate treatment not related to adolescent pregnancy such as nutrition information) rather than merely assessing outcomes in the control or comparison group.

A final related moderator of interest differentiated tier 1 programs with previous evidence of effectiveness, thus testing replications, from tier 2 new or previously untested programs.

### Individual Participant Data Collection

Through the HHS Office of Adolescent Health, we requested de-identified individual participant data (IPD) for each eligible study identifying each participant's group assignment, demographic characteristics, baseline sexual risk and behavior, sexual risk and behavior at follow-up, and study design variables. Once data were received, a doctoral-level researcher reviewed each data set for data consistency and completeness, with follow-up communication with grantees to resolve any questions or identified inconsistencies.

### Analysis

We conducted all analyses in Stata/MP, version 14.2 (StataCorp LP, College Station, TX). Using data extracted from the evaluation reports, we first conducted meta-analyses to estimate overall effects on the various primary outcomes. Because many studies provided multiple effect sizes, we used robust variance estimation to handle the associated statistical dependencies, using the *robmeta* package.<sup>9–12</sup> We used the intercept from unconditional random effects robust variance estimation meta-regression models to estimate the mean effect sizes. We examined heterogeneity by using  $\tau^2$  as the measure of between-study effect size variability and  $I^2$  as

**TABLE 1—Characteristics of the Studies, Intervention Programs, and Youth Participants: Meta-Analysis of Federally Funded Adolescent Pregnancy Prevention Program Evaluations, United States, 2010–2015**

Characteristic	No. (%) or Mean $\pm$ SD (Range)	Valid No.
<b>Methodological characteristics of the included studies</b>		
<b>Design</b>		
Randomized controlled trial (individual)	20 (45)	44
Randomized controlled trial (cluster)	20 (45)	44
Controlled quasi-experimental design	4 (9)	44
Active comparison condition	16 (36)	44
<b>Attrition</b>		
Attrition at first follow-up	0.25 $\pm$ 0.17 (0.03–0.62)	44
Attrition at last follow-up	0.39 $\pm$ 0.19 (0.12–0.88)	24
Differential attrition <sup>a</sup>	0.04 $\pm$ 0.03 (0.00–0.26)	206
<b>Posttest assessment timing<sup>a</sup></b>		
0 < X $\leq$ 3 mo	30 (13)	226
3 < X $\leq$ 6 mo	72 (32)	226
6 < X $\leq$ 9 mo	75 (33)	226
9 < X $\leq$ 12 mo	6 (3)	226
12 < X mo	43 (19)	226
<b>Characteristics of the intervention programs</b>		
Evidence-based program	20 (45)	44
<b>Primary program focus</b>		
Abstinence only	1 (2)	44
Sexual health	26 (59)	44
Youth development	13 (30)	44
HIV/AIDS prevention	3 (7)	44
Reproductive health services	1 (2)	44
<b>Program components<sup>b</sup></b>		
Condom demonstrations	16 (36)	44
Service learning	6 (14)	44
Role plays	27 (61)	44
Games	9 (20)	44
Reflective exercises	11 (25)	44
Mentoring/tutoring	1 (2)	44
Individualized counseling	2 (5)	44
Direct provision of health services	4 (9)	44
Parent activities	9 (20)	44
Community outreach	0 (0)	44
Positive role model	7 (16)	44
<b>Group size</b>		
Individualized	5 (11)	44
Small groups (< 10)	6 (14)	44
Large groups	29 (66)	44
Online	3 (7)	44
Other (mixed individual/group)	1 (2)	44
<b>Program setting</b>		
Classroom	24 (55)	44
Health clinic	4 (9)	44
Community	10 (23)	44
Other setting	6 (14)	44

*Continued*

TABLE 1—Continued

Characteristic	No. (%) or Mean $\pm$ SD (Range)	Valid No.
<b>Group composition<sup>c</sup></b>		
Program delivered to single-gender groups	10 (23)	44
Program delivered to mixed-gender groups	34 (77)	44
<b>Gender-specific programming<sup>d</sup></b>		
Gender-specific program (i.e., girls only)	8 (18)	44
Mixed-gender program	36 (82)	44
<b>Program delivery personnel</b>		
Medical professionals	1 (2)	44
Health educators	16 (36)	44
Classroom teachers	7 (16)	44
Peer educators	2 (5)	44
Mixed	7 (16)	44
Other	11(25)	44
<b>Program contact: frequency</b>		
Daily	4 (10)	39
3–4 times per wk	5 (13)	39
1–2 times per wk	24 (62)	39
Less than weekly	6 (15)	39
<b>Program contact: duration</b>		
	36.51 $\pm$ 47.79 (1–189)	
< 4 wk	8 (20)	40
4–11.9 wk	10 (25)	40
12–51.9 wk	15 (38)	40
$\geq$ 52 wk	7 (18)	40
<b>Program contact: hours</b>		
	62.52 $\pm$ 180.95 (2–1134)	
< 10 h	15 (34)	44
10–29.9 h	18 (41)	44
30–49.9 h	6 (14)	44
$\geq$ 50 h	5 (11)	44
<b>Fidelity to program protocol</b>		
	0.95 $\pm$ 0.06 (0.69–1.00)	42
<b>Participant engagement</b>		
Attendance	0.81 $\pm$ 0.17 (0.35–1.00)	42
Retention	0.77 $\pm$ 0.21 (0.20–1.00)	41
<b>Characteristics of youth participants</b>		
Percentage boys	0.38 $\pm$ 0.19 (0–0.75)	44
Percentage Black	0.38 $\pm$ 0.33 (0.01–1)	39
Percentage Hispanic	0.37 $\pm$ 0.29 (0.03–1)	38
Percentage White	0.23 $\pm$ 0.21 (0–0.83)	35
Average age, y	14.42 $\pm$ 2.14 (11–19)	44
Proportion recent sexual activity at baseline	0.31 $\pm$ 0.30 (0–1)	27
Proportion recent unprotected sexual activity at baseline	0.14 $\pm$ 0.19 (0–0.74)	31
Proportion control group sexually active at posttest	0.38 $\pm$ 0.30 (0–1)	44

Note. Means and standard deviations shown for continuous measures; frequencies and percentages shown for dichotomous measures.

<sup>a</sup>Estimates calculated at effect size level ( $n = 226$ ). Intervals measured from the end of the intervention as reported by study authors.

<sup>b</sup>Program components are not mutually exclusive, and therefore the sum is more than 100%.

<sup>c</sup>Some programs deliver services to mixed-gender groups while others separate boys and girls into single-gender groups. Programs that are designed for a single gender are classified in the latter category.

<sup>d</sup>Certain programs are designed to be provided exclusively to members of a single gender; in our sample all such programs were girl-only.

the proportion of total variability attributable to between-study differences.

Then, synthesizing all effect sizes regardless of outcome category, we estimated a series of robust variance estimation mixed-effects meta-regression models to examine associations between effect sizes and the different candidate moderators. To aid in interpretation, we report the results from an omnibus F-test assessing the overall significance of each moderator block.<sup>13,14</sup>

Finally, we used IPD to further examine variability in program effects across the participant characteristics of age, gender, and race/ethnicity. Not all authors provided IPD, so the final 1-stage IPD analysis model included a mixture of IPD and aggregate data, estimated by using the MEQRLOGIT command in Stata.<sup>15–17</sup>

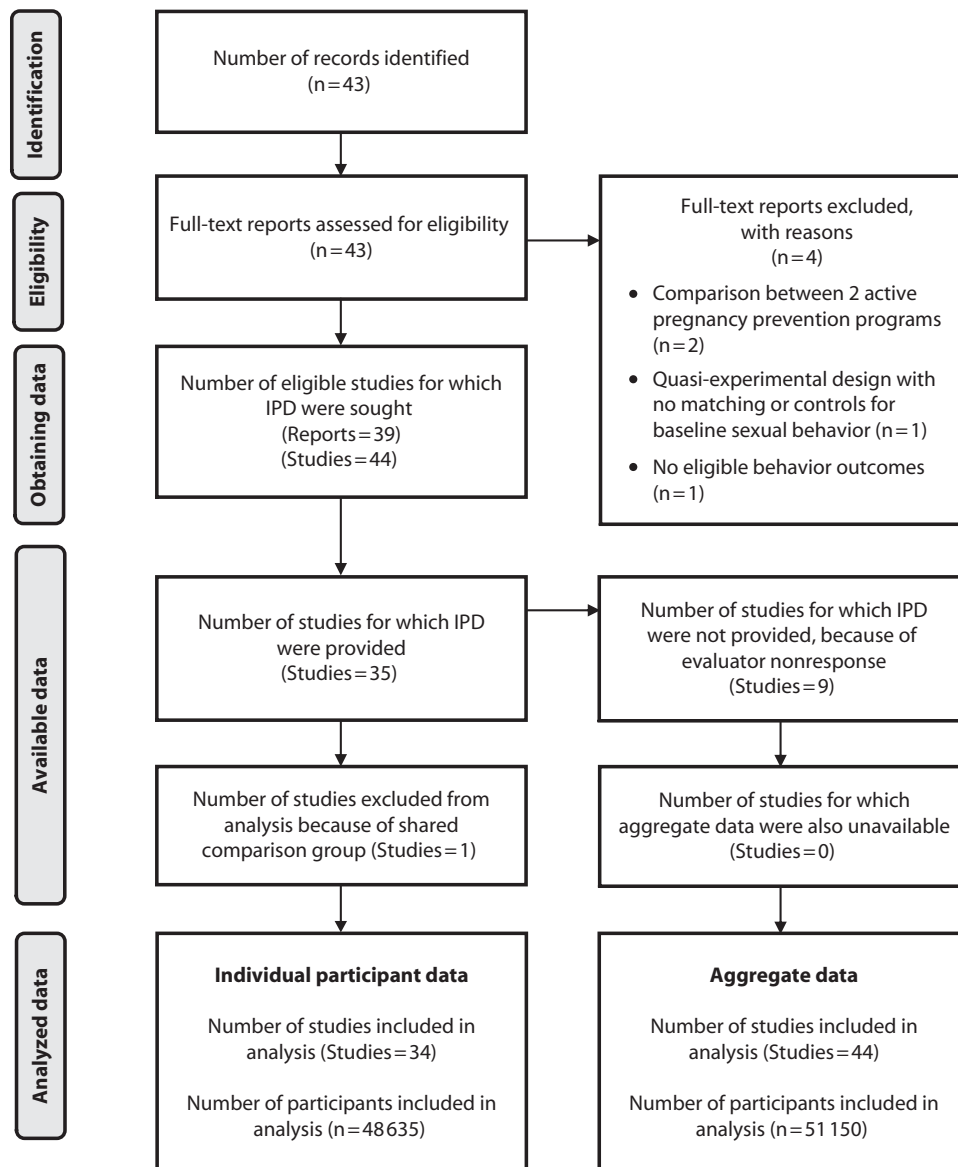
## RESULTS

We deemed 39 of 43 candidate reports from HHS eligible (Figure 1) and presented findings from 44 independent samples, comprising 51 150 youths.<sup>18–56</sup> We received IPD for 35 of the 44 independent study samples (80%), and included 34 of these (48 635 youths) in the final IPD meta-analysis.

### Study Characteristics

Table 1 summarizes the studies' methodological characteristics, the characteristics of the evaluated programs, and the demographic composition of samples. Overall study quality was high, with most (90%) of the 44 studies using a randomized design. Effect sizes were reported at varying durations of follow-up. Many studies reported effects at multiple follow-ups (Appendix A) although few of those follow-ups were more than 9 months after the end of program participation. Approximately half of the programs were replications of evidence-based program models (45%). Program implementation, attendance, and retention data indicated that most programs were delivered with high fidelity, although some struggled with participant engagement.

The demographic composition of samples varied considerably. Among the studies that



Note. IPD = individual participant data.

**FIGURE 1—Study Identification Flow Diagram for Meta-Analysis of US Federally Funded Adolescent Pregnancy Prevention Program Evaluations**

reported youths' sexual activity at baseline, the average proportion of youths recently sexually active was 31%, and the proportion who reported recent unprotected sex was 14%.

### Overall Program Effects

In each of the 9 outcome categories, we found small but statistically insignificant mean effect sizes generally favoring the programs. For most outcomes we found little evidence of variation in effect sizes

across studies, evidenced by the small  $\tau^2$  estimates and generally small  $I^2$  estimates (Table 2).

### Moderators of Program Effects

We found little evidence of moderator effects related to program or participant characteristics (Table 3). Nor was there a statistically significant difference between the mean effect sizes for tier 1 programs with previous evidence of effectiveness and the new or untested tier 2 programs.

We found no statistically reliable evidence of associations between effect sizes and any of the moderators related to program elements and implementation features with 2 exceptions (Table 3). There was a modest association between gender-specific (girl-only) programs and effect sizes ( $b = 0.11$ ; 95% CI =  $-0.01, 0.23$ ). Of the 44 studies, 8 evaluated programs designed specifically for girls; the 51 effect sizes from those 8 studies yielded a statistically significant average effect size indicating favorable program effects on sexual behavior (LOR = 0.11; 95% CI = 0.01, 0.20;

**TABLE 2—Overall Effects of Adolescent Pregnancy Prevention Programs by Outcome: Meta-Analysis of Federally Funded Adolescent Pregnancy Prevention Program Evaluations, United States, 2010–2015**

Outcome Category	No. of Studies ( <i>k</i> )	No. of Effect Sizes Reported	Effect Size, LOR or Hedges' <i>g</i> (95% CI)	$\tau^2$	$I^2$ , %
Ever had sex	24	43	0.04 (–0.04, 0.13)	0.01	24
Recent sexual activity	25	53	0.02 (–0.05, 0.09)	0.01	25
Recent unprotected sexual activity	35	87	0.06 (–0.03, 0.14)	0.02	27
Proportion of recent sexual experiences that were unprotected <sup>a</sup>	1	1	–0.29 (–0.85, 0.27)	NA	NA
Sexually transmitted infections	2	2	0.17 (–0.70, 1.04)	0.28	70
Ever pregnant or parent	7	15	0.11 (–0.31, 0.53)	0.07	51
Recent pregnancy or parenting	5	9	0.10 (–0.77, 0.96)	0.00	0
No. of sexual partners	5	9	0.03 (–0.04, 0.09)	0.00	0
No. of sexual experiences	1	1	0.03 (–0.12, 0.18)	NA	NA

Note. CI = confidence interval; LOR = log odds ratio; NA = not applicable.

<sup>a</sup>Without condoms or other forms of birth control.

OR = 1.12). This result should be interpreted cautiously given the insignificant difference between such programs and programs designed for mixed genders ( $P = .07$ ). We also found a modest relationship between participant group size and effect size: programs with individualized delivery (the reference category) had larger effect sizes than those with larger groups. However, only the specific comparison with the small groups category was statistically significant, with programs delivering services in small group settings less effective than individualized programs ( $b = -0.20$ ; 95% CI =  $-0.05, -0.34$ ).

We found no statistically significant associations between program effects and any of the demographic characteristics measured at the aggregate study level (Table 3) or those measured at the individual level in any of the outcome categories for which individual participant data were consistently available: ever had sex, recent sexual activity, recent unprotected sexual activity, and pregnancy or parenting (see Appendix A, Figures A–D, for detailed individual-level findings).

None of the study method moderator variables related to risk of bias (i.e., assignment to conditions, overall and differential attrition) or type of control or comparison group were significantly associated with effect sizes.

## DISCUSSION

Although several evaluations in this sample reported evidence of effects, we found little

evidence of associations between those impacts and the isolated program and participant characteristics that could be examined in this meta-analysis. The results do suggest that programs designed exclusively for girls may be distinctively effective. The reason for this relationship is unclear—our analysis of IPD found no evidence that adolescent pregnancy prevention programs were more effective for girls in general, and the 8 studies in this subgroup shared other characteristics such as serving older, higher-risk adolescents in individualized settings, the latter feature being the only other program characteristic associated with positive effects.

## Interpretation of Findings

The primary aim of this meta-analysis was to identify any characteristics of the adolescent pregnancy prevention programs in the HHS-funded sample associated with positive effects on sexual risk behavior and pregnancy. Such information could potentially guide selection, implementation, and improvement of these programs for adolescent participants. Despite examining a range of plausibly influential program characteristics, this endeavor was not very successful. The modest variation around small mean effects on the behavioral outcomes addressed by these programs revealed little differential effectiveness to be explained by the candidate moderator variables.

One possible explanation for this result is the relatively low risk profile of the participant

samples, most of whom were young (14.4 years on average at baseline) and not sexually active (on average, only 38% of youths in control groups were sexually active at follow-up). Most outcomes were measured within 9 months of program participation in this sample, which may have been too soon for prevention programs to show effects on behaviors that had not yet fully emerged. Many of the studies did find impacts on such intermediate outcomes as knowledge, attitudes, and skills (not reported here). Those results are promising, but there is not yet strong evidence linking changes in these outcomes with changes in later risky sexual behavior. The field may need to identify better predictors of later risky sexual behavior as well as develop and validate more proximal outcome measures for younger youths.

## Limitations

As we synthesized evidence only from US programs federally funded by HHS, the findings may not be broadly generalizable (e.g., to adolescents in other countries or to programs not funded by that agency).

Another limitation is that the sample of studies included in the meta-analysis was not sufficiently large to allow us to test combinations of factors that might collectively make up an effective program (e.g., individualized gender-specific programming delivered to older adolescents).

**TABLE 3—Relationships Between Program-Level Moderators and Effect Sizes: Meta-Analysis of Federally Funded Adolescent Pregnancy Prevention Program Evaluations, United States, 2010–2015**

Moderator	b (95% CI)
<b>Prior evidence (program tier; F = 1.66; P = .21)</b>	
New or previously untested program	0 (Ref)
Evidence-based program	-0.07 (-0.20, 0.05)
<b>Program focus (F = 0.23; P = .80)</b>	
Sexual health	0 (Ref)
Youth development	0.04 (-0.12, 0.20)
Other	0.06 (-0.28, 0.41)
<b>Program components (F = 0.80; P = .63)</b>	
Condom demonstrations	0.10 (-0.06, 0.27)
Service learning	0.08 (-0.35, 0.50)
Role plays	-0.02 (-0.25, 0.21)
Games	0.15 (-0.11, 0.41)
Reflective exercises	0.05 (-0.10, 0.21)
Direct provision of health services	0.20 (-0.13, 0.54)
Parent activities	-0.03 (-0.24, 0.19)
Positive role model	-0.03 (-0.27, 0.21)
<b>Group size (F = 2.54; P = .20)</b>	
Individualized	0 (Ref)
Small groups (< 10)	-0.20* (-0.34, -0.05)
Large groups	-0.10 (-0.24, 0.05)
Other (mixed individual/group)	-0.07 (-0.26, 0.12)
<b>Group composition (F = 0.88; P = .37)</b>	
Mixed-gender delivery	0 (Ref)
Same-gender delivery	0.06 (-0.08, 0.19)
<b>Gender-specific programming: girls only (F = 4.78; P = .07)</b>	
0.11 (-0.01, 0.23)	
<b>Program contact (F = 0.11; P = .90)</b>	
At least weekly contact	-0.03 (-0.15, 0.10)
Contact hours	0.00 (-0.00, 0.00)
<b>Program setting (F = 1.83; P = .21)</b>	
Classroom	0 (Ref)
Community	0.10 (-0.10, 0.31)
Other	0.10 (-0.02, 0.22)
<b>Program delivery personnel (F = 0.30; P = .75)</b>	
Health educators	-0.05 (-0.24, 0.14)
Classroom teachers	0 (Ref)
Other	-0.00 (-0.20, 0.19)

*Continued***TABLE 3—Continued**

Moderator	b (95% CI)
<b>Fidelity and participant engagement (F = 0.73; P = .56)</b>	
Fidelity	-0.04 (-1.58, 1.49)
Mean attendance	0.64 (-2.86, 4.14)
Mean retention	-0.22 (-2.71, 2.27)
<b>Participant characteristics (F = 0.60; P = .70)</b>	
Percentage boys	-0.18 (-0.60, 0.25)
Percentage Black	-0.00 (-0.41, 0.41)
Percentage Hispanic	-0.04 (-0.39, 0.31)
Average age	-0.04 (-0.14, 0.07)
Risk (control event rate)	0.21 (-0.47, 0.89)
<b>Study method (F = 0.91; P = .52)</b>	
Randomized controlled trial	0.50 (-0.06, 1.06)
Overall attrition	0.17 (-0.39, 0.72)
Differential attrition	2.56 (-0.55, 5.67)
Active control group	-0.03 (-0.19, 0.13)

Note. CI = confidence interval. All meta-regression models estimated by using robust variance estimation to handle statistically dependent effect sizes. The analytic sample size was  $n = 44$  studies and 226 effect sizes for all analyses.

## Public Health Implications

Though the findings of this meta-analysis offer limited guidance for policymakers, the evidence suggesting that individualized programs and those designed specifically for girls may be distinctively effective offers some direction for program developers and implementers. Because no programs in this sample targeted only boys, we could not examine the effectiveness of comparable programs for them. Nevertheless, it seems worthwhile to reexamine assumptions about the effectiveness of the numerous interventions for coed classrooms in changing the behavioral outcomes such programs are intended to prevent. *AJPH*

## CONTRIBUTORS

All authors contributed substantially to conceptualizing the study, interpreting the data, revising the article, and reviewing and approving the final version of the article. In addition, R. Juras and E. Tanner-Smith analyzed the data and drafted the article's content.

## ACKNOWLEDGMENTS

This work was funded under contract HHSP2332015000691 from the US Department of Health and Human Services.

Many faculty and staff at Abt Associates and Vanderbilt University contributed to this study. We would especially like to thank Alicia Hymel, Heather Hensman Kettrey,

Jennifer Norvell, Leah Roberts, and Emily Farnen at Vanderbilt and Yeqin He at Abt Associates for their outstanding data coding work, and Austin Nichols at Abt Associates for his helpful comments on drafts of the article.

**Note.** The Department of Health and Human Services was given the opportunity to review and comment on a draft of this article, but not to approve or reject publication based on content. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Health and Human Services.

## CONFLICTS OF INTEREST

Three of the study's authors (R. J., M. K., and J. L.) are co-authors on 3 of the primary reports included in the meta-analysis. Therefore, these authors were not involved in screening or data extraction for those reports; external and independent data collectors were used to extract all such data.

## HUMAN PARTICIPANT PROTECTION

No protocol approval was necessary for this study because no human participants were involved.

## REFERENCES

- Lugo-Gil J, Lee A, Vohra D, Adamek K, Laco J, Goesling B. Updated findings from the HHS Teen Pregnancy Prevention Evidence Review: July 2014 through August 2015. Washington, DC: US Department of Health and Human Services; 2016.
- Kappeler E. Building the evidence to prevent adolescent pregnancy. *Am J Public Health*. 2016;106(suppl 1):S5.
- Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*. 2009;6(7):e1000097.
- Juras R, Tanner-Smith E, Kelsey M, Lipsey M, Layzer J. Design and analysis plan quantitative synthesis of federally-funded teen pregnancy prevention programs. HHS Contract HHSP2332015000691. May 2, 2016. Available at: <https://www.hhs.gov/ash/oah/sites/default/files/analysis-plan-050216-final-508.pdf>. Accessed January 31, 2019.
- Office of Adolescent Health, Department of Health and Human Services. Grantee evaluations FY2010–2014. 2016. Available at: <https://www.hhs.gov/ash/oah/evaluation-and-research/grantee-led-evaluation/2010-2014-grantees/index.html>. Accessed June 13, 2018.
- Lipsey MW, Wilson DB. *Practical Meta-analysis*. Thousand Oaks, CA: Sage Publications; 2001.
- Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *J Educ Behav Stat*. 1981; 6(2):107–128.
- Higgins JPT, Deeks JJ, Altman DG. Special topics in statistics. In: Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: John Wiley & Sons; 2008:481–529.
- Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Res Synth Methods*. 2010;1(1):39–65.
- Tanner-Smith EE, Tipton E. Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Res Synth Methods*. 2014;5(1):13–30.
- Tipton E. Robust variance estimation in meta-regression with binary dependent effects. *Res Synth Methods*. 2013;4(2):169–187.

12. Tipton E. Small simple adjustments for robust variance estimation with meta-regression. *Psychol Methods*. 2015; 20(3):375–393.
13. Pustejovsky JE. clubSandwich: cluster-robust (sandwich) variance estimators with small-sample corrections. R package version 0.0.0.9000. Available at: <https://github.com/jepusto/clubSandwich>. Accessed June 13, 2018.
14. Tipton E, Pustejovsky JE. Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *J Educ Behav Stat*. 2015; 40(6):604–634.
15. Fisher DJ, Copas AJ, Tierney JF, Parmar MK. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. *J Clin Epidemiol*. 2011;64(9):949–967.
16. Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med*. 2008;27(11):1870–1893.
17. Riley RD, Steyerberg EW. Meta-analysis of a binary outcome using individual participant data and aggregate data. *Res Synth Methods*. 2010;1(1):2–19.
18. Abe Y, Barker LT, Chan V, Eucogco J. Early findings from the evaluation of the Pono Choices Program—a culturally-responsive teen pregnancy and sexually transmitted infection prevention program for middle school youth in Hawai'i. Columbia, MD: IMPAQ International; 2016.
19. *¡Cúdate!*: interim impact report. Cambridge, MA: Abt Associates; 2016.
20. Reducing the risk: short-term impact report. Cambridge, MA: Abt Associates; 2016.
21. Safer sex intervention: short-term impact report. Cambridge, MA: Abt Associates; 2016.
22. Evaluation of Will Power/Won't Power in Los Angeles County. Los Angeles, CA: Advanced Empirical Solutions; 2015.
23. Calise TV, Chow W, Dore KF. Evaluation of Healthy Futures in public middle schools in three northeastern Massachusetts cities: findings from an innovative teen pregnancy prevention program. Boston, MA: JSI Research and Training Institute; 2015.
24. Carter SL, Beadnell B, Vanslyke J. Evaluation of the Web of Life teen pregnancy prevention program: findings from an innovative positive youth development approach for American Indian youth. Albuquerque, NM: National Indian Youth Leadership Project; 2015.
25. Covington RD, Goesling G, Trenholm C, et al. Interim impacts of the AIM 4 Teen Moms program. Princeton, NJ: Mathematica Policy Research; 2015.
26. Coyle KK, Potter SC, Glassman JR, McDade-Montez L, Unti T. Evaluation of It's Your Game . . . Keep It Real in South Carolina: final report. Scotts Valley, CA: ETR Associates; 2015.
27. Coyle K, Anderson P, Laris BA, Unti T, Franks H, Glassman J. Evaluation of It's Your Game . . . Keep It Real in Houston, TX: final report. Scotts Valley, CA: ETR Associates; 2016.
28. Crean HF, Seibold-Simpson SM, Jambon M, Kreipe RE. Evaluation of the Teen Outreach Program® in Rochester, New York: findings from the replication of an evidence-based teen pregnancy prevention program. Rochester, NY: University of Rochester School of Nursing; 2016.
29. Cunningham MR, van Zyl MA, Borders KW. Evaluation of Love Notes and Reducing the Risk in Louisville, Kentucky. Louisville, KY: University of Louisville Research Foundation; 2016.
30. Daley EM, Buhi ER, Wang W, et al. Evaluation of Wyman's Teen Outreach Program® in Florida: final impact report for Florida Department of Health. Findings from the replication of an evidence-based teen pregnancy prevention program. Tallahassee, FL: Florida Department of Health; 2015.
31. Dierschke N, Gelfond J, Lowe D, Schenken RS, Plastino K. Evaluation of Need to Know (N2K) in South Texas: findings from an innovative demonstration program for teen pregnancy prevention program. San Antonio, TX: The University of Texas Health Science Center at San Antonio; 2015.
32. Eichner J, Salaway J, Smith-Jones J, McCall R. Evaluation of Seventeen Days in Ohio, Pennsylvania, and West Virginia: findings from the replication of an evidence-based teen pregnancy prevention program. Pittsburgh, PA: University of Pittsburgh, Office of Child Development; 2015.
33. Francis K, Woodford M, Kelsey M. Evaluation of the Teen Outreach Program in Hennepin County, MN: findings from the replication of an evidence-based teen pregnancy prevention program. Cambridge, MA: Abt Associates; 2015.
34. Goesling B, Colman S, Scott M, Cook E. Impacts of an enhanced family health and sexuality module of the HealthTeacher middle school curriculum. Princeton, NJ: Mathematica Policy Research; 2014.
35. Herrling S. Evaluation of the Children's Aid Society (CAS)-Carrera Adolescent Pregnancy Prevention Program in Chicago, IL: findings from the replication of an evidence-based teen pregnancy prevention program. Accord, NY: Philliber Research and Evaluation; 2016.
36. Kissinger P, Schmidt N, Green J, et al. Evaluation of BUtiful: an Internet pregnancy prevention for older teenage girls in New Orleans, Louisiana. New Orleans, LA: Tulane University School of Public Health and Tropical Medicine; 2015.
37. LaChausse R. Evaluation of the Positive Prevention PLUS teen pregnancy prevention program. San Bernardino, CA: California State University; 2015.
38. Martin S, Hill A, Nye M, Hollman-Billmeier K. Evaluation of Alaska Promoting Health Among Teens, Comprehensive Abstinence and Safer Sex (AKPHAT) in Alaska. Anchorage, AK: Institute of Social and Economic Research, University of Alaska Anchorage; 2015.
39. Philliber AE, Philliber S, Brown S. Evaluation of the Teen Outreach Program® in the Pacific Northwest. Accord, NY: Philliber Research and Evaluation; 2016.
40. Philliber AE, Philliber S. Evaluation of the Teen Outreach Program® in Kansas City, Missouri. Accord, NY: Philliber Research and Evaluation; 2016.
41. Piotrowski ZH, Hedeker D. Evaluation of the Be The Exception sixth-grade program in rural communities to delay the onset of sexual behavior. *Am J Public Health*. 2016;106(suppl 1):S132–S139.
42. Piotrowski ZH, Hedeker D. Evaluation of Positive Potential middle school program in rural Northwest Indiana communities. Chicago, IL: University of Chicago; 2015.
43. Piotrowski ZH, Hedeker D. Evaluation of the Positive Potential Be The Exception grade 6 program in predominantly rural communities. Chicago, IL: University of Chicago; 2015.
44. Jenner E, Walsh S, Jenner LW, Demby H, Gregory A, Davis E. Evaluation of Becoming a Responsible Teen: findings from the replication of an evidence-based teen pregnancy prevention program. New Orleans, LA: The Policy and Research Group; 2015.
45. Robinson WT, Kaufman R, Cahill L. Evaluation of the Teen Outreach Program in Louisiana. New Orleans, LA: Louisiana State University Health Sciences Center at New Orleans, School of Public Health; 2016.
46. Rotz D, Goesling B, Crofton M, Trenholm C, Manlove J, Welti K. Interim impacts of Teen PEP in New Jersey and North Carolina high schools. Cambridge, MA: Mathematica Policy Research; 2016.
47. Ruwe MB, McCloskey L, Meyers A, Prudent N, Foureau-Dorsinville MK. Evaluation of Haitian-American Responsible Teen: findings from the replication of an evidence-based teen pregnancy prevention program in Eastern Massachusetts. Boston, MA: Boston Medical Center; 2016.
48. Schwinn T, Kaufman CE, Black K, et al. Evaluation of Circle of Life in tribes of the Northern Plains: findings from an innovative teen pregnancy prevention program. Denver, CO: University of Colorado Denver, Centers for American Indian and Alaska Native Health; 2015.
49. Seshadri R, Smithgall C, George R, et al. Evaluation of Teen Outreach Program in Chicago: final impact report for Chicago public schools. Chicago, IL: Chapin Hall at the University of Chicago; 2015.
50. Slater HM, Mitschke DB. Evaluation of the Crossroads Program in Arlington, TX: findings from an innovative teen pregnancy prevention program. Arlington, TX: University of Texas at Arlington; 2015.
51. Smith KV, Dye C, Cook E, Rosinsky K, Scott M. Interim impacts of the Gender Matters Program. Princeton, NJ: Mathematica Policy Research; 2016.
52. Smith K, Rotz D, Goesling B, Cook E, Murphy K, Stevens J. Interim impacts of the Teen Options to Prevent Pregnancy program. Princeton, NJ: Mathematica Policy Research; 2015.
53. Tucker T. Evaluation of the Carrera Program: findings from the replication of an evidence-based teen pregnancy prevention program. Atlanta, GA: Tressa Tucker and Associates; 2015.
54. Usera JJ, Curtis KM. Evaluation of the Ateyapi Identity Mentoring Program in South Dakota: findings from the replication of an evidence-based teen pregnancy prevention program. Sturgis, SD: Delta Evaluation Consulting LLC; 2015.
55. Vyas A, Wood S, Landry M, Douglass G, Fallon S. The evaluation of the Be Yourself / Sé Tú Mismo in Montgomery and Prince George's Counties, Maryland. Washington, DC: The George Washington University Milken Institute School of Public Health; 2015.
56. Walker EM, Inoa R, Coppola N. Evaluation of Promoting Health Among Teens! abstinence-only intervention in Yonkers, NY. Princeton, NJ: Sametric Research; 2016.