

assistance: outcomes and lessons learned. *Health Promot Pract.* 2009;10(1):38S–44S.

5. Zief SG, Knab J, Cole RP. A framework for evaluation technical assistance.

*Am J Public Health.* 2016;106(suppl 1):S24–S26.

6. Knab J, Cole RP, Zief SG. Challenges and lessons learned from providing large-scale evaluation technical assistance

to build the adolescent pregnancy evidence base. *Am J Public Health.* 2016;(suppl 1):S26–S28.

7. Margolis AL, Roper A. Practical experience from the Office of

Adolescent Health's large scale implementation of an evidence-based teen pregnancy prevention program. *J Adolesc Health.* 2014;54:S10–S14.

# A Framework for Evaluation Technical Assistance

The primary goal of the evaluation technical assistance (TA) was to help the grantees and their evaluators produce rigorous evidence of program effectiveness that would meet the US Department of Health and Human Services (HHS) Teen Pregnancy Prevention (TPP) evidence review standards, described in Cole et al.<sup>1</sup> To achieve this goal, the evaluation TA team created a framework for TA that included an initial design review, ongoing monitoring and group TA, and a series of evaluation deliverables designed to serve as stepping stones toward the final evaluation report.

This editorial is the second in a series of related opinion pieces. As noted above, the first editorial described the context of the evaluation TA contract, and outlined the HHS evidence standards that were used as an operational goal for a high-quality evaluation. This second editorial describes the TA framework. The final editorial provides a summary of the challenges and lessons learned from the evaluation TA project, and offers insight into the ways that evaluation TA has been improved for future grantee funded evaluations.<sup>2</sup>

## DESIGN REVIEWS (YEAR 1)

The first step for funded grantees was receiving approval of their proposed designs. We

reviewed the planned evaluation designs to determine whether they could potentially meet the HHS TPP evidence review standards, provided the evaluations were well implemented over the subsequent five years, and be sufficiently powered to detect statistically significant impacts. Unlike the evidence review, which assessed completed studies, the design review assessed evaluation *plans* to determine the likelihood that aspects of the design and implementation would not introduce sample bias and the final evidence would meet the evidence review standards. We took the evidence review standards and worked backward to be sure the proposed designs and data collection plans were aligned with the final standards.

The team used a design review template to systematically review each proposed design, assessing features including, but not limited to the (random) assignment procedure, the approach for recruiting and consenting sample members, and plans for sample retention and data collection. These features were assessed to determine whether they were potential threats to the validity of the study. For instance, when we encountered scenarios that might compromise the integrity of the random assignment design (for example, moving youths from the condition to which they were randomly assigned to a different condition), we worked to the

extent possible with the grantee and evaluator to reduce potential threats to the study meeting standards. In addition, depending on feasibility and logistical and resource constraints, we attempted to improve the statistical power of each study design, if needed, to increase the likelihood of the study showing statistically significant impacts.

During the design review phase, the evaluation TA team provided feedback to grantees through written reviews, telephone calls, and in-person discussions. The Office of Adolescent Health (OAH) decided when to approve each design after receiving input from the evaluation TA team, at which point the evaluation TA team shifted gears toward monitoring the quality of the ongoing evaluations.

## ONGOING MONITORING AND GROUP TA (YEARS 2–4)

The second phase of evaluation TA was monitoring the implementation of the evaluation and identifying solutions to real-world problems that affect nearly all rigorous impact

evaluations once they go into the field. The evaluation TA team conducted ongoing monitoring through individual TA and also provided group TA for common issues faced across the evaluations.

- **Ongoing monitoring:** Each grantee and their evaluator worked closely with a TA liaison, who served as the point person for the broader evaluation TA team. In conjunction with the project officer, the TA liaison held monthly monitoring calls during the implementation period to identify risks to the integrity of the study and to brainstorm suggested solutions for addressing them. Individual TA was an important component of the TA framework, so the TA could be customized to the needs and experiences of the evaluation team.
- **Group TA:** Above and beyond the individual TA provided to each grantee and their evaluators, the evaluation TA team provided a wide array of group TA. The group TA included presentations of timely evaluation topics at grantee conferences, webinars and conference calls, and written briefs on evaluation topics, such as descriptions of the evidence review standards, methods for handling missing or inconsistent self-report

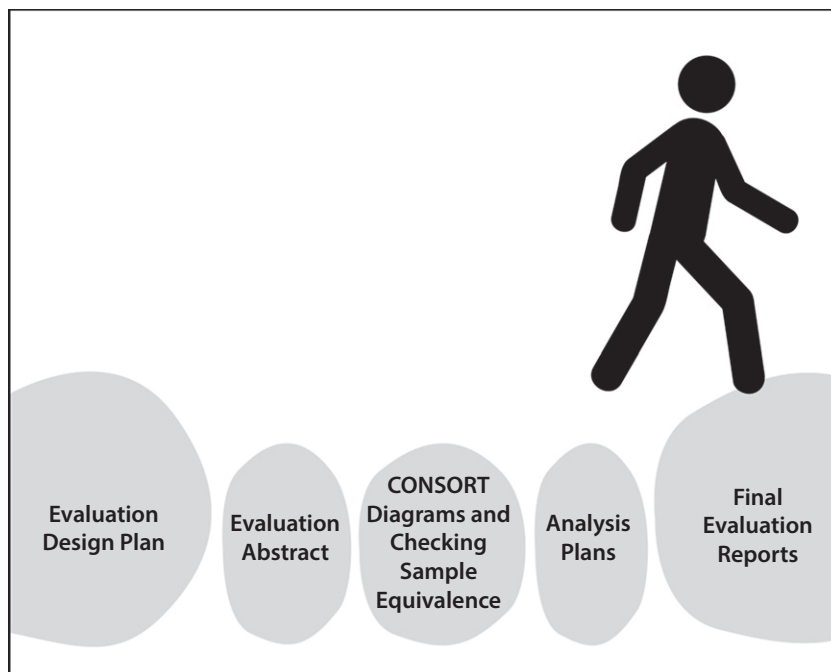
## ABOUT THE AUTHORS

Susan Goerlich Zief and Russell P. Cole are Senior Researchers at Mathematica Policy Research, Princeton, NJ. Jean Knab is an Associate Director of Human Services Research at Mathematica Policy Research.

Correspondence should be sent to Jean Knab, Mathematica Policy Research, PO Box 2393, Princeton, NJ 08543-2393 (e-mail: jknab@mathematica-mpr.com). Reprints can be ordered at <http://www.ajph.org> by clicking the "Reprints" link.

This editorial was accepted June 23, 2016.

doi: 10.2105/AJPH.2016.303365



**FIGURE 1—Products that Serve as Stepping Stones Toward the Final Evaluation Report for the Office of Adolescent Health.**

survey responses, methods for adjusting for non-independence in clustered designs, and ways to create a well-matched sample in a quasi-experimental design.

## STEPPING STONES (YEARS 3–4)

Once the evaluations were well under way, OAH asked the evaluation TA team to identify a series of products that would serve as the stepping stones toward a final evaluation report that would clearly describe the study and its findings, and provide sufficient documentation for the HHS TPP evidence review team. To help grantees and evaluators prepare for this final evaluation report, the evaluation TA team requested three interim products. The evaluation TA team created a template for each product and OAH

required grantees to submit the products as part of the cooperative agreements. The evaluation TA team reviewed and commented on each product before it could be incorporated into the final evaluation reports. (Figure 1).

- **Evaluation abstract:**

Grantees and their evaluators completed a structured abstract to serve as an executive summary of the planned evaluation. The preliminary abstracts were published before the evaluations were completed, serving as a means to disseminate information about the forthcoming studies as well as to informally “register” the studies. The abstracts would eventually become the executive summary to each final report.

- **CONSORT diagrams and checking sample equivalence:** Twice a year, the grantee and their evaluator submitted CONSORT

diagrams to document retention of the eligible evaluation sample. They also assessed the equivalence of the treatment and control groups for the sample enrolled and completing surveys to date. The TA liaisons used the information to assess study progress against the HHS evidence standards and to provide targeted TA. For example, TA liaisons calculated sample attrition from the CONSORT diagrams, using sample sizes at random assignment and each follow-up survey and provided suggestions for boosting response rates, as needed. If the sample was found to have high attrition or lacked equivalence, the liaison provided TA on conducting an analysis that could meet HHS evidence standards. The CONSORT diagram would ultimately be used to

document sample flow in the final report and accompanying journal articles, and the final baseline equivalence assessment was included in the final evaluation report.

- **Analysis plans:** Federal agencies, such as HHS and the US Department of Education, and scientific journals increasingly require evaluations to prespecify their planned impact analyses. These plans identify the outcomes of interest and the analytic strategies before all data have been collected and analyzed. Vetting a plan at an early stage of an evaluation supports a more transparent process and provides the funder and an evaluator with a road map for the final report. OAH required that their grantees and evaluators submit an impact analysis plan, following a template developed by the evaluation TA team. The impact analysis plan template requested information on how variables would be coded, methods for handling missing and inconsistent data, approaches for estimating impacts, and sensitivity analyses to assess the robustness of findings. In addition, the impact analysis plan template was structured so that several sections of the final plan could be directly incorporated into the final reports, reducing the burden on the grantees and evaluators at the end of the evaluation. OAH requested that the final, approved impact analysis plan also be submitted to <https://clinicaltrials.gov>. OAH also requested an implementation analysis plan to understand how the data on attendance, implementation fidelity, and quality would be used.

## CONCLUSIONS

The evaluation TA contract funded by OAH was successful in establishing a framework for a large-scale evaluation TA effort. The team developed processes and report templates that facilitated training and TA to more than 100 grantees and evaluator partners, as well as timely reporting to

OAH. The framework that was developed, with improvements made from lessons learned,<sup>2</sup> is also being used for OAH's cohort 2 and has been adopted by similar efforts outside of OAH. *AJPH*

*Susan Goerlich Zief, PhD*  
*Jean Knab, PhD*  
*Russell P. Cole, PhD*

# Challenges and Lessons Learned From Providing Large-Scale Evaluation Technical Assistance to Build the Adolescent Pregnancy Evidence Base

This is the third editorial in a series of related opinion pieces. The first editorial provided a description of the Office of Adolescent Health's (OAH) investment in teen pregnancy prevention (TPP) programs and the evaluation technical assistance (TA) contract that was intended to support funded grantees.<sup>1</sup> That editorial also outlined the US Department of Health and Human Services (HHS) evidence standards that guided the evaluation TA contract. The second editorial detailed the activities conducted under the evaluation TA contract with the first cohort of funded grantees.<sup>2</sup> Implementing a large-scale TA effort was not without challenges. These challenges were necessary and expected, as the federal government looks to support large-scale grantee-led evaluations to build the evidence base in TPP. The lessons learned from these challenges laid the groundwork for a second round of evaluation TA with a second cohort of TPP grantees that began in summer 2015.

These challenges and their solutions can also inform other federally sponsored evaluation TA efforts. The following is a discussion of the major challenges in providing TA to the grantees in cohort 1. These challenges were identified through the TA activities laid out in Zief et al.,<sup>2</sup> primarily our monitoring calls and document reviews.

## CHALLENGE 1: NO STANDARDIZED APPROACH

Federal efforts to assess impact evaluations for the credibility of their causal inferences began with the Department of Education's What Works Clearinghouse in 2004 and have spread throughout other agencies and offices, such as HHS and the Department of Labor. These agencies have each developed similar sets of standards that are used to systematically and consistently assess the internal validity of completed studies in

## CONTRIBUTORS

All authors contributed equally to this editorial.

## ACKNOWLEDGMENTS

This work was conducted under a contract (HHSP233201300416G) with the Office of Adolescent Health within the Department of Health and Human Services (HHS).

## REFERENCES

1. Cole RP, Zief SG, Knab J. Establishing an evaluation technical assistance contract

to support studies in meeting the HHS evidence standards. *Am J Public Health*. 2016;106(suppl 1):S22–S24.

2. Knab J, Cole RP, Zief SG. Challenges and lessons learned from providing large-scale evaluation technical assistance to build the adolescent pregnancy evidence base. *Am J Public Health*. 2016; 106(suppl 1):S26–S28.

products that could inform the broader field and future grantees and evaluators. The contract produced research briefs on planning evaluations designed to meet evidence standards, coping with missing data and clustering in randomized controlled trials, baseline equivalence and matching, attrition in randomized controlled trials, best practices for school and district recruitment, and a primer on the evidence standards. The dissemination materials, available on the OAH Web site,<sup>3</sup> supplied a road map for the second cohort of grantees that did not exist for the first cohort.

## CHALLENGE 2: NOT THE SOLE BENCHMARK

The HHS evidence review assesses the internal validity of the impact findings to document the extent to which the results are credible and could therefore guide policy and program decisions. However, for most policymakers and researchers, the key question of interest is

their field. But no publicly available road maps existed within HHS, or any other federal agency, for designing and implementing an evaluation that has a good chance of resulting in a completed study that meets evidence review standards. Because many grantees and evaluators were initially unaware or only partially aware of the standards, they were not considering the evidence standards when designing their evaluations. So this contract focused on building a process (the framework noted in Zief et al.<sup>2</sup>) for training the grantees on the evidence standards and best practices for meeting those standards. This was accomplished through providing formal and informal training during the TA process and developing written

## ABOUT THE AUTHORS

Jean Knab is an Associate Director of Human Services Research at Mathematica Policy Research, Princeton, NJ. Russell P. Cole and Susan Goerlich Zief are Senior Researchers at Mathematica Policy Research.

Correspondence should be sent to Jean Knab, Mathematica Policy Research, PO Box 2393, Princeton, NJ 08543-2393 (e-mail: jknab@mathematica-mpr.com). Reprints can be ordered at <http://www.ajph.org> by clicking the "Reprints" link.

This editorial was accepted June 23, 2016.

doi: 10.2105/AJPH.2016.303358