



ICGFM DC Forum Luncheon

FOREIGN ASSISTANCE:

**Agencies Can Improve the Quality and Dissemination of
Program Evaluations
(GAO-17-316)**

April 5, 2017

Contents

- Background
- Objectives
- Scope and Methodology
- Quality Assessment
- Cost Analysis
- Dissemination Review
- Conclusions and Recommendations
- Agency Response

Background

- The U.S. government plans to spend approximately \$35 billion on foreign assistance in 2017.
- Evaluation is an essential tool for U.S. agencies to assess and improve the results of their programs.
- The 2010 GPRA Modernization Act, the 2010 Presidential Policy Directive on Global Development policy, and the 2016 Foreign Aid Transparency and Accountability Act (FATAA) call for an increased focus and guidelines for conducting foreign assistance evaluations.
- Prior GAO reports have identified challenges in the design, implementation, and dissemination of foreign assistance evaluations.

Objectives

This report examines:

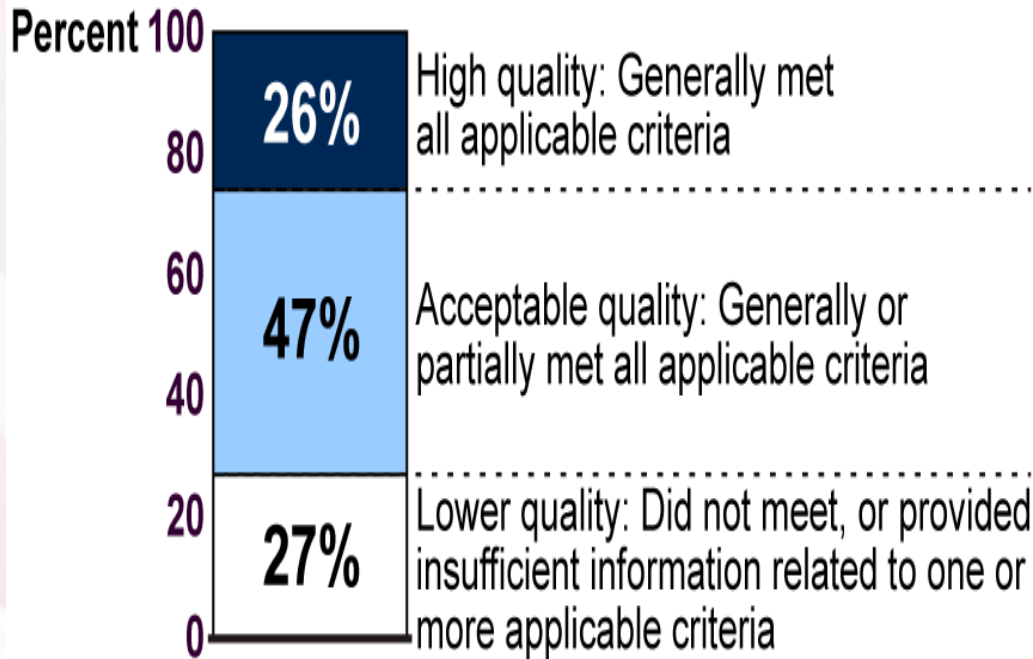
- (1) the extent to which foreign assistance program evaluations met key evaluation **quality** criteria,
- (2) the **costs** of the agencies' evaluations and factors that affect these costs, and
- (3) the extent to which the agencies ensure the **dissemination** of evaluation reports within the agency and to the public.

Scope and Methodology

- Scope:
 - Focused on the six agencies providing the largest amounts of U.S. foreign aid: USAID, State, MCC, HHS, USDA, and DOD.
 - At HHS, DOD, and USDA, we reviewed selected programs: HHS's Centers for Disease Control and Prevention programs under the President's Emergency Plan for AIDS Relief, USDA's Foreign Agricultural Service food assistance, and DOD's Global Train and Equip program
 - For our evaluation quality assessment, we selected a representative sample of 173 fiscal year 2015 completed evaluations for foreign assistance programs. Evaluations reviewed included all State, MCC, DOD-GTE, and USAID net impact evaluations, and a random sample of the remaining evaluations by agency.
 - For our cost assessment, we collected cost information for 76 final evaluations by MCC, State, USAID and USDA.
- Methodology:
 - Assessed the selected evaluations against leading evaluation quality criteria;
 - Analyzed final evaluations' cost and contract documents; and
 - Reviewed agency websites and dissemination procedures.

Quality

Most Foreign Aid Evaluations Had High or Acceptable Quality



Quality criteria

Design:

1. Questions align with intervention goals
2. Indicators/measures appropriate
3. Design appropriate

Implementation:

4. Target population and sampling appropriate
5. Data collection appropriate
6. Data analysis appropriate

Conclusions:

7. Conclusions supported
8. Recommendations and lessons learned justified

Source: GAO analysis of agency data. | GAO-17-316

Note: The confidence intervals for our estimates of the quality of agency evaluations according to these categories did not exceed ± 8 percent.

Quality

Overall Quality Varied by Agency

	Percentage of evaluations						
	All agencies	DOD	HHS	MCC	State	USAID	USDA
High quality: Generally met all applicable criteria.	26	0	35	44	4	26	21
Acceptable quality but could be improved: Generally or partially met all applicable criteria but did not generally meet all.	47	50	38	44	48	49	48
Lower quality: Did not meet, or provided insufficient information related to, one or more applicable criteria.	27	50	26	13	48	26	30
Number of evaluations reviewed	173	4	34	16	23	63	33

Legend: DOD = Department of Defense Global Train and Equip program; HHS = Department of Health and Human Services Centers for Disease Control & Prevention—President’s Emergency Plan for AIDS Relief; MCC = Millennium Challenge Corporation; State = Department of State; USAID = U.S. Agency for International Development; USDA = U.S. Department of Agriculture Foreign Agricultural Service food assistance programs.

Source: GAO analysis of fiscal year 2015 agency evaluation reports.

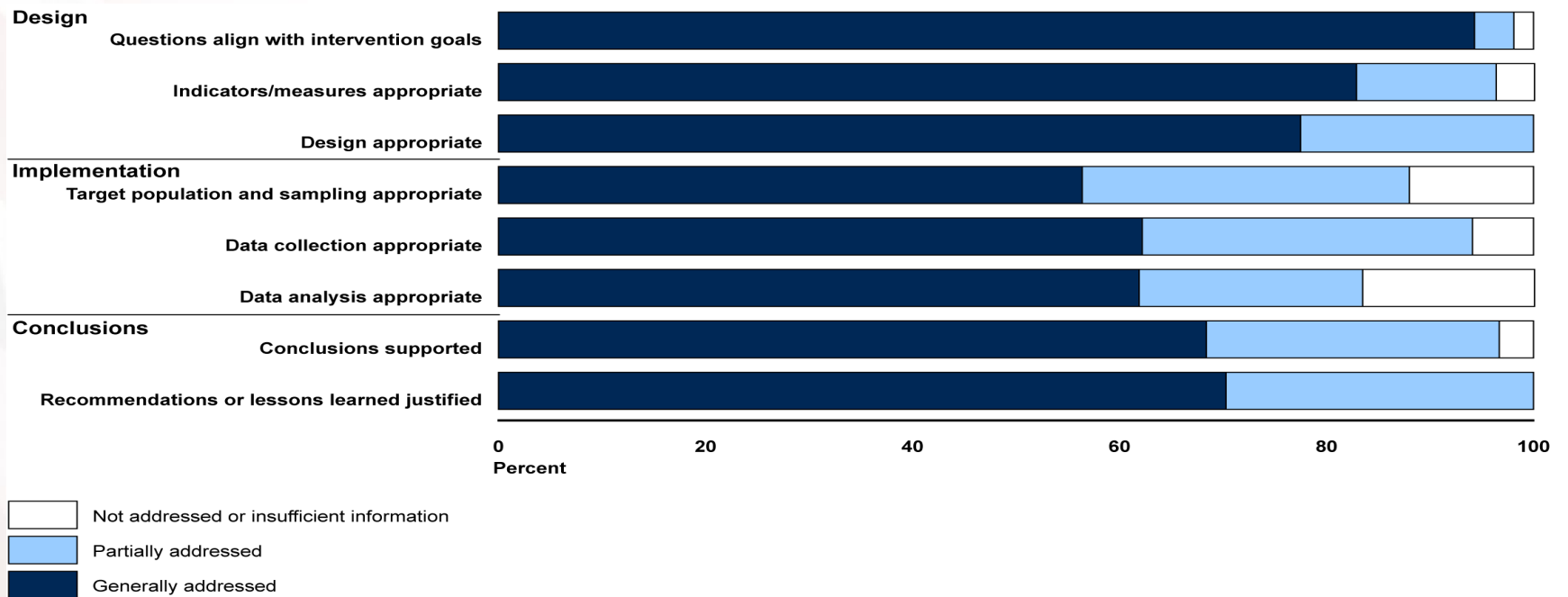
Notes: Percentages for all agencies combined and for USAID are weighted to reflect the evaluations in the population that were not selected for the sample. The confidence interval for all six agencies did not exceed ± 8 percent for the seven criteria that were based on the full sample of 173 evaluations and did not exceed ± 9 percent for the variable that relied on the 161 evaluations that had recommendations and lessons learned. The confidence intervals for the estimates for USAID, USDA, and HHS did not exceed ± 11 percent except for the variable that relied on the 161 evaluations that had recommendations and lessons learned, where it did not exceed ± 14 percent.

We assessed 161 evaluations against all eight quality criteria. We assessed the 12 evaluations that did not include recommendations or lessons learned against the remaining seven applicable criteria.

Columns may not sum to 100 percent because of rounding.

Quality

Evaluation Implementation Posed the Greatest Challenges



Source: GAO analysis of fiscal year 2015 agency evaluation reports. | GAO-17-316

Notes: We assessed 161 evaluations against all eight quality criteria. We assessed 12 evaluations that did not include recommendations or lessons learned against the remaining seven applicable criteria.

Percentages for all agencies combined and for the U.S. Agency for International Development (USAID) are weighted to reflect the evaluations in the population that were not selected for the sample. The confidence interval for all six agencies did not exceed ±8 percent for the seven criteria that were based on the full sample of 173 evaluations and did not exceed ±9 percent for the variable that relied on the 161 evaluations that had recommendations and lessons learned.

Percentages shown for each criterion may not sum to 100 because of rounding.

We reviewed evaluations of foreign assistance programs administered by the Department of Defense Global Train and Equip program; Department of Health and Human Services's Centers for Disease Control & Prevention—President's Emergency Plan for AIDS Relief; Millennium Challenge Corporation; Department of State; USAID; and U.S. Department of Agriculture Foreign Agricultural Service food assistance programs.



Quality

Assessment by Quality Criteria Varied by Agency

Criterion	Percentage of evaluations that generally met quality criteria						
	Total	DOD	HHS	MCC	State	USAID	USDA
1. Study questions align with the key stated goal(s) of the intervention.	94	100	85	94	96	96	94
2. The chosen indicators/measures are appropriate for the study objectives.	83	50	97	88	48	83	88
3. The evaluation design is appropriate given the study questions.	78	25	79	88	57	80	73
4. The target population and sampling for the evaluation are appropriate for the study questions.	56	0	62	63	43	59	48
5. The data collection is appropriate for the study questions.	62	0	76	69	35	63	61
6. The data analysis appears appropriate to the task.	62	25	74	69	48	63	52
7. Conclusions are supported by the available evidence.	68	25	65	75	61	73	52
8. Recommendations and lessons learned are supported by the available evidence.	70	33	80	79	86	65	82
Number of evaluations	173	4	34	16	23	63	33

Source: GAO analysis of fiscal year 2015 agency evaluations. GAO-17-316

Notes: We assessed 161 evaluations with recommendations against all eight quality criteria and assessed 12 evaluations without recommendations against the remaining seven applicable criteria.

Percentages for all agencies combined and for USAID are weighted to reflect the evaluations in the population that were not selected for the sample. The confidence interval for all six agencies did not exceed ±8 percent for the seven criteria that were based on the full sample of 173 evaluations and did not exceed ±9 percent for the variable that relied on the 161 evaluations that had recommendations and lessons learned. The confidence intervals for the estimates for USAID, USDA, and HHS did not exceed ±11 percent except for the variable that relied on the 161 evaluations that had recommendations and lessons learned, where it did not exceed 14 percent.



Quality Evaluation Design

Criterion	Response percentage				Number	Maximum Confidence Interval
	Generally addressed	Partially addressed	Insufficient information	Not at all		
1. Study questions align with the key stated goal(s) of the intervention.	94	4	2	0	173	+/-3.2 percent
2. The chosen indicators/measures are appropriate for the study objectives.	83	13	3	0	173	+/-5.5 percent
3. The evaluation design is appropriate given the study questions.	78	22	0	0	173	+/-5.9 percent
Does the evaluation assess any net impacts?	Yes: 18	No: 81	0		173	+/-0.4 percent
Does the evaluation state baselines?	Yes: 51	No: 47	2		170	+/-7.2 percent
Does the evaluation state specific targets?	Yes: 54	No: 46	1		170	+/-6.9 percent
Does the evaluation assess processes such as program implementation?	Yes: 88	No: 12	0		173	+/-4.2 percent
Does the evaluation establish criteria such as established plans, budgets, timeframes, and targets?	Yes: 52	No: 48	0		149	+/-7.8 percent
Is the study performed by a third party evaluator?	Yes: 80	No: 13	6		173	+/-3.8 percent
Are potential conflicts of interest discussed?	Yes: 38	No: 62	0		173	+/-7.1 percent

Source: GAO analysis of fiscal year 2015 agency evaluations. GAO-17-316

Notes: Rows may not add up to 100 percent due to rounding. Percentages are weighted to reflect the evaluations in the population that were not selected for the sample.



Quality Evaluation Implementation

Criterion	Response percentage				Number	Maximum Confidence Interval
	Generally addressed	Partially addressed	Insufficient information	Not at all		
4. The target population and sampling for the evaluation are appropriate for the study questions.	56	32	11	1	173	+/-7.3 percent
Does the sampling frame appear appropriate?	53	22	24	1	164	+/-7.5 percent
Is random sampling used?	Yes: 34	No: 44	19	Not appl. 3	173	+/-6.1 percent
Are the (random) sampling methods appropriate?	82	7	11	0	69	+/-10.2 percent
Is nonrandom sampling used?	Yes: 70	No: 7	21	Not appl. 2	173	+/-6.3 percent
Are the sampling methods for nonrandom sampling appropriate?	51	25	23	2	114	+/-9.1 percent
5. The data collection is appropriate for the study questions.	62	32	6	0	173	+/-7.1 percent
Are the data collection methods specified for each question?	75	15	6	4	173	+/-6.4 percent
Do data collection procedures appear to ensure the reliability of the data?	40	24	24	13	173	+/-7.0 percent
6. The data analysis appears appropriate to the task.	62	22	16	0	173	+/-7.1 percent
Are data analysis methods clearly specified for each question?	61	15	9	15	173	+/-7.2 percent

Source: GAO analysis of fiscal year 2015 agency evaluations. GAO-17-316

Notes: Rows may not add up to 100 percent due to rounding. Percentages are weighted to reflect the evaluations in the population that were not selected for the sample.

Quality Conclusions of Evaluations

Criterion	Response percentage				Number	Maximum Confidence Interval
	Generally addressed	Partially addressed	Insufficient information	Not at all		
7. Conclusions are supported by the available evidence.	68	28	3	0	173	+/-6.6 percent
Are the main study limitations identified (in design, data collection methods, and analysis)?	56	29	0	15	173	+/-7.1 percent
Does the evaluation consider the possibility of unintended side effects of the intervention(s)?	Yes: 28	No: 72	0		173	+/-6.4 percent
8a. Recommendations are supported by the available evidence.	74	26	0	0	149	+/-7.3 percent
8b. Lessons learned are supported by the available evidence.	83	17	0	0	119	+/-6.8 percent

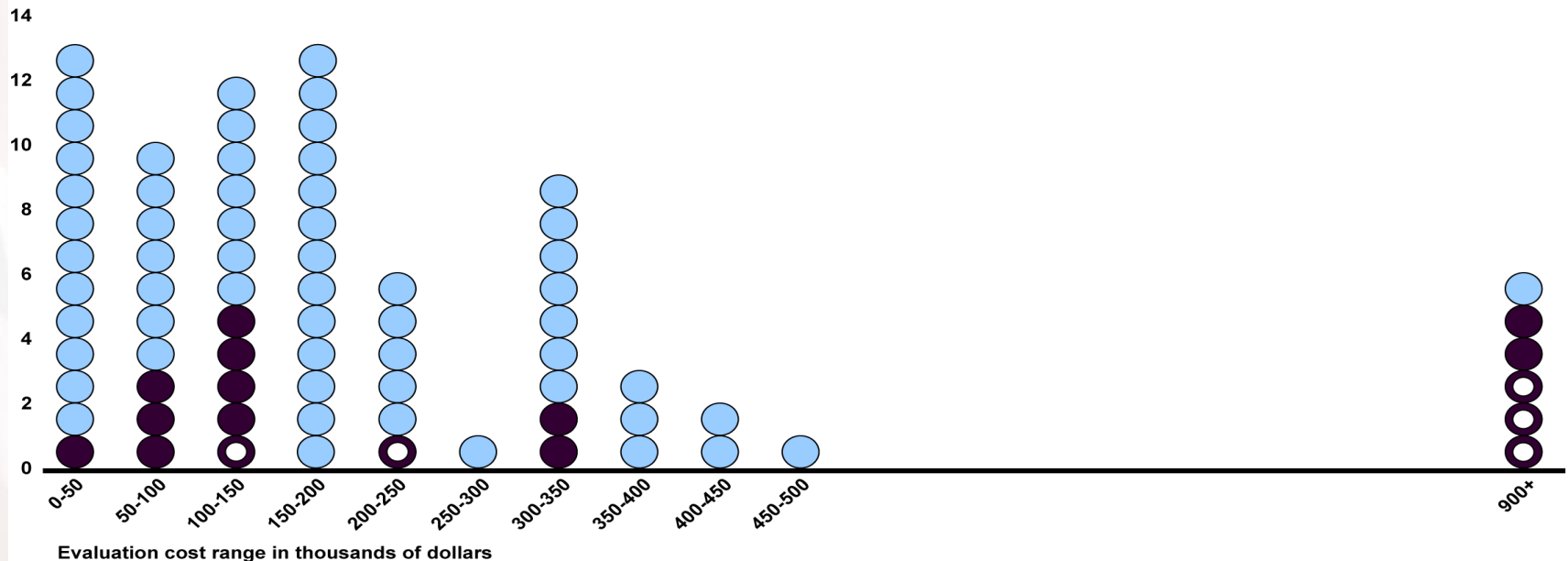
Source: GAO analysis of fiscal year 2015 agency evaluations. GAO-17-316

Notes: Rows may not add up to 100 percent due to rounding. Percentages are weighted to reflect the evaluations in the population that were not selected for the sample.

Cost

Costs Ranged Widely, but The Majority Were Less Than \$200,000

Number of evaluations in cost range



- Performance evaluation
- Net impact evaluation
- Randomized controlled trial

Source: GAO analysis of agency data. | GAO-17-316

Note: Data include costs for 76 MCC, State, USAID, and USDA evaluations. We could not obtain specific cost information for the Department of Defense's Global Train and Equip and the Department of Health and Human Services' President's Emergency Plan for AIDS Relief evaluations because they were not separately tracked by the agencies, contractors, or implementing partners.

Cost

Costs Varied by Agency and Evaluation Type

Agency/ evaluation type	Number of evaluations	Evaluation cost (dollars in thousands)			
		Average	Median	Minimum	Maximum
MCC	12	\$520.2	268.9	9.6	2,230.4
Net impact	6	\$863.8	657.1	85.0	2,230.4
Performance	6	\$176.5	107.5	9.6	457.0
State	16	\$248.6	177.7	38.3	902.1
Net impact	2	\$211.0	211.0	79.0	343.0
Performance	14	\$254.0	177.7	38.3	902.1
USAID	38	\$207.8	147.3	22.1	1,012.5
Net impact	8	\$340.2	117.2	97.1	1,012.5
Performance	30	\$172.5	159.8	22.1	407.5
USDA	10	\$129.9	87.9	26.5	401.2
Net impact	1	\$36.1	36.1	36.1	36.1
Performance	9	\$140.3	122.8	26.5	401.2

Sources: GAO analysis of evaluation contracts, invoices, and related documents; data from Federal Procurement Data System–Next Generation. | GAO-17-316.

Note: We could not obtain specific cost information for the Department of Defense’s Global Train and Equip and the Department of Health and Human Services’ President’s Emergency Plan for AIDS Relief evaluations because they were not separately tracked by the agencies, contractors, or implementing partners.

Cost

High-Quality Evaluations Tend to Cost More

	Number of evaluations	Cost (dollars in thousands)			
		Average	Median	Minimum	Maximum
High quality	15	\$543.3	307.4	97.1	2,230.4
Acceptable quality	39	\$198.8	169.6	9.6	902.1
Lower quality	22	\$159.6	98.8	25.9	981.8

Sources: GAO analysis of foreign assistance evaluation reports, evaluation contracts, invoices, and related documents as well as data from Federal Procurement Data System–Next Generation. GAO-17-316

Note: High-quality evaluations generally met all applicable quality criteria. Acceptable-quality evaluations partially or generally met all applicable quality criteria. Lower-quality evaluations did not meet, or provided insufficient information related to, one or more quality criteria.

- Some lower-cost evaluations also met all quality criteria.
- We found the following factors that influence evaluation costs:
 - data collection methods,
 - frequency of data collection,
 - evaluation duration, and
 - evaluation design.

Dissemination Evaluations Are Generally Available Online

Effective dissemination practices ^a	DOD	HHS	MCC	State	USAID	USDA
1. Generally post nonsensitive evaluations online	N/A ^b	Yes	Yes	Yes	Yes	No ^d
2. Provide a search engine that can find the evaluations	N/A	Yes ^c	Yes	Yes	Yes	N/A ^d
3. Post evaluations in the timeframe required by the agency	N/A	No	No	N/A ^e	Yes	N/A
4. Make sensitive evaluations accessible internally	Yes	N/A ^f	N/A ^f	Yes	Yes	N/A ^f
5. Require planning for the dissemination of evaluations	Yes ^g	Yes	Yes	No	Yes	Yes
6. Use means other than public posting to disseminate evaluations	Yes ^g	Yes	Yes	Yes	Yes	Yes

Source: GAO analysis of agency evaluation guidance, evaluation and dissemination documents, and websites. GAO-17-316

^aWe identified six dissemination practices, based on federal and other guidance, that agencies should generally use to ensure the effective dissemination of evaluation reports.

^bThe four DOD evaluations we reviewed, for the Global Train and Equip (GT&E) program, were designated "sensitive" and thus were not required to be posted on a public website.

^cThe Centers for Disease Control and Prevention (CDC) is in the process of adding all evaluations from fiscal year 2015 to the CDC Stacks website. This website has a search engine that can be used to locate individual evaluations. CDC reported that 49 of the 51 evaluations from fiscal year 2015 had been posted on the CDC Stacks website as of December 2016.

^dUSDA does not post evaluations online. According to USDA officials, the department is in the process of developing procedures for making nonsensitive evaluations public.

^eIn fiscal year 2015, State did not have a policy requiring that evaluations be posted within a certain timeframe. In fiscal year 2016, State revised its guidance to require that evaluations be posted online 90 days after completion.

^fHHS, MCC, and USDA did not designate as sensitive any of the evaluations we reviewed.

^gDOD plans for evaluation dissemination by identifying potential users of the evaluation on a standard e-mail distribution list. DOD uses this standard e-mail distribution list to disseminate GT&E program evaluations via e-mail to congressional stakeholders, as required by law, as well as internal stakeholders.

Dissemination

Timeliness and Planning Can Be Improved

- *Public posting.* Most agencies posted nonsensitive reports on a public website, which is searchable with agency search engines. All agencies have websites to make sensitive evaluations available internally.
 - USDA has not developed procedures for reviewing and preparing its evaluations for public posting.
- *Timeliness.* Some HHS reports and more than half of MCC reports were posted a year or more after completion. PEPFAR guidance requires that evaluations be posted within 90 days of completion, while HHS requires that evaluations be publicly posted within a year of their completion.
- *Dissemination planning.* USAID included dissemination plans in most evaluations, and other agencies will require such plans in the future.
 - State does not currently have a policy requiring a plan that identifies potential users and the means of dissemination.

Conclusions

- Over 70 percent of the evaluations we reviewed (124 out of 173) were of high or acceptable quality.
 - The greatest areas for improvement were in evaluation implementation-sampling, data collection and analysis.
- Over 90 percent (70 out of 76) of the evaluations whose costs we obtained cost less than \$500,000. The remaining 10% or less of the evaluations cost over \$900,000.
 - 5 out of the 6 were net impact evaluations; 3 out of the 6 used RCTs. Evaluation design, frequency of primary data collection and duration of the evaluation affect costs. MCC's evaluations had the highest median cost, half of MCC's evaluations were net impact evaluations.
- Agencies generally post their evaluations to their websites. Timeliness of posting and dissemination planning can be improved.

Recommendations — GAO recommends that:

- Each agency develop a plan for improving the quality of evaluations for the programs included in our review, focusing on areas where our analysis has shown the largest areas for potential improvement.
- HHS-CDC update its guidance and practices on the posting of evaluations to require PEPFAR evaluations to be posted within the timeframe required by PEPFAR guidance.
- MCC adjust its evaluation practices to make evaluation reports available within the timeframe required by MCC guidance.
- State amend its evaluation policy to require the completion of dissemination plans for all agency evaluations.
- USDA implement guidance and procedures for making FAS evaluations available online and searchable on a single website that can be accessed by the general public.

Agency Comments and Our Evaluation

- The agencies generally concurred with our recommendations.
- State, USAID, and USDA each identified specific steps they plan to take to implement the recommendations.
- DOD and MCC have not yet identified any specific steps.
- HHS is taking action to address our dissemination recommendation but did not comment on our recommendation that it develop a plan for improving the quality of evaluations.
- GAO will follow up with each agency to confirm steps taken to address our recommendations.
 - Agencies are required to outline plans to implement GAO recommendations within 60 days.



GAO on the Web

Connect with GAO on [LinkedIn](#), [Facebook](#), [Flickr](#), [Twitter](#), [YouTube](#) and our Web site: <http://www.gao.gov/>
Subscribe to our [RSS Feeds](#) or [E-mail Updates](#). Listen to our [Podcasts](#) and read [The Watchblog](#)

Congressional Relations

Katherine Siggerud, Managing Director, siggerudk@gao.gov
(202) 512-4400, U.S. Government Accountability Office
441 G Street, NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov
(202) 512-4800, U.S. Government Accountability Office
441 G Street, NW, Room 7149, Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov
(202) 512-4707, U.S. Government Accountability Office,
441 G Street NW, Room 7814, Washington, DC 20548

Copyright

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.
