



PEPFAR

U.S. President's Emergency Plan for AIDS Relief

Evaluation Standards of Practice

Version 3.0

12/15/2017

U.S. Department of State
Office of the U.S. Global AIDS Coordinator and Health Diplomacy

PEPFAR EVALUATION STANDARDS OF PRACTICE

TABLE OF CONTENTS

Contents

Introduction.....	1
Section I – Terms and Key Concepts.....	3
Section II – PEPFAR Evaluation Standards of Practice.....	7
Section III – Planning, Implementation, Reporting, and Review.....	14
Section IV – Roles and Responsibilities.....	17

Tables

Table 1. Stakeholder roles and responsibilities.....	3
Table 2. Individual roles and responsibilities.....	18
Table 3. Individual roles and responsibilities.....	20

Figures

Figure 1. Planning, implementation, reporting, and review process.....	14
--	----

Appendices

Appendix A – Definition of Terms.....	A1
Appendix B – Required Components of an Evaluation Report.....	B1
Appendix C – Evaluation Resources.....	C1
Appendix D – Reference Materials.....	D1

ACRONYMS

AM	Activity Manager
AOR	Agreement Officer Representative
APR	Annual Program Results
COI	Conflict of Interest
COP	Country Operational Plan
COR	Contracting Officer Representative
ESoP	Evaluation Standards of Practice
GAO	Government Accountability Office
HQ	Headquarters
IAD	Implementing Agency Designee
IE	Impact Evaluation
IOM	Institute of Medicine
IS	Implementation Science
IRB	Institutional Review Board
M&E	Monitoring and Evaluation
MER	Monitoring, Evaluation, and Reporting
S/GAC	State Department/Office of the U.S. Global AIDS Coordinator
OU	Operating Unit
OR	Operations Research
PEPFAR	President's Emergency Plan for AIDS Relief
PHE	Public Health Evaluation
PM	Project Manager
POC	Point of Contact
ROP	Regional Operational Plan
SI	Strategic Information
SOW	Scope of Work
USG	United States Government

Preface

The impetus for the development of the Evaluation Standards of Practice (ESoP) was the growing need within PEPFAR to maximize the utility of evaluations and, more significantly, to respond to a call for improved evaluations and transparency from the Government Accounting Office (GAO) and the Institute of Medicine (IOM).

The inaugural version of the PEPFAR ESoP was published in January 2014. It contained the 11 standards of practice that should be followed to ensure high standards of evaluation planning, implementation, dissemination, and use. The ESoP served as a guide to improve the quality of evaluations and reinforce the role of using evaluation findings in programmatic decision-making.

The second version of ESoP vs. 2 was published in September 2015. This version included guidance for operationalizing planning, implementation, and reporting of PEPFAR evaluations. In ESoP vs. 2, new sections were added pertaining to operational issues for planning, implementation, reporting, and review, roles and responsibilities, and tools.

ESoP vs. 3 has been expanded to include implementation science and operations research, and more clearly delineates definitions for evaluation and research types. The tools to support application of this guidance are available on pepfar.net. The key tools found at the site include: the evaluation inventory and the adherence checklist. The supporting tools for the inventory include the data dictionary which defines each field in the inventory and a template for collecting relevant information in preparation for DATIM entry.

We understand that changes to evaluation reporting will take additional effort. We will make all efforts to provide as much guidance and orientation to these reporting requirements, but we believe that this updated ESoP vs. 3 and the accompanying revised tools and materials will improve the quality of evaluation and research reporting and will enable PEPFAR to have better access to these data for program planning and decision making.

For further questions, please contact your evaluation Points of Contact at your Operating Unit and/or Agency. The below email is available for any questions for S/GAC.

SGAC_EWG@state.gov

Introduction

The President's Emergency Plan for AIDS Relief (PEPFAR) has been at the forefront of implementing robust program monitoring to track progress towards reaching epidemic control. In addition to routine program data and surveillance, evaluation and research play a key role in generating evidence needed to know what works and how to implement efficiently and cost-effective interventions to achieve epidemic control. Evaluation is useful to assess implementation fidelity, to demonstrate efficacy and effectiveness, to measure intervention success at changing behavior, and to understand cost implications of one approach vs. another. Furthermore, implementation science and operations research are useful to study how evidence-based programs can best be applied in real world settings; and to improve the quality of service delivery while achieving public health scale. To that end, ESoP will enable PEPFAR to collect data on what activities are occurring and what questions are being asked, where they are being conducted, and how findings can be leveraged to inform programmatic decision-making across PEPFAR programs in a timely way. ESoP generates an inventory of PEPFAR funded evaluations, helps to assess their adherence to the ESoP, and provides the needed evidence to work towards achievement of the highest standards in evaluation practice.

ESoP was developed, in part, as a response to the Government Accountability Office (GAO) and the Institute of Medicine (IOM) recommendations on how to improve evaluation quality in PEPFAR. To reach the 90-90-90 goals and reach epidemic control, there is a critical need for high quality evaluations driven by evidence gaps, that are implemented in line with quality standards, and for which results are disseminated publicly and used by relevant stakeholders^{1,2}. This is consistent with recommendations from GAO³ and IOM⁴.

The purpose of ESoP vs. 3 is to reinforce the standards of evaluation practice deemed most relevant to conducting quality evaluations within PEPFAR and to inform improved decision-making. Many of these evaluation standards of practice are promoted through international and professional evaluation associations,^{5,6} and to a large extent, are already integrated into PEPFAR implementing agency policies and strategy documents^{7,8,9}. In addition to articulating the standards of practice, this document also contains guidance to assist in the implementation of these standards and the associated requirements. These sections offer an overview of PEPFAR evaluation planning, implementation, reporting, and review, as well as guidance on the roles and responsibilities of the relevant stakeholders and individuals who will implement, oversee, and report on the process.

Anticipated users of this document include evaluators, researchers, persons who commission and procure evaluation and research services, providers of technical assistance for evaluations and research, and host country partners, among others.

The goal of the ESoP is to improve evaluation and research planning, implementation, oversight, and quality

¹ Government Accountability Office (GAO), President's Emergency Plan for AIDS Relief: Agencies Can Enhance Evaluation Quality, Planning, and Dissemination, GAO-12-673, May 31, 2012.

² Institute of Medicine (IOM), 2013. Evaluation of PEPFAR, The National Academies Press: Washington, DC.

³ GAO, 2012, *op. cit.*

⁴ IOM, 2013, *op. cit.*

⁵ American Evaluation Association (AEA), An Evaluation Roadmap for a More Effective Government, September 2010. <http://www.eval.org/EPTF/aea10.roadmap.101910.pdf>

⁶ African Evaluation Association, African Evaluation Guidelines - Standards and Norms, <http://www.afrea.org>

⁷ USAID, USAID Evaluation Policy, Learning from Experience. <http://transition.usaid.gov/evaluation/USAIDEvaluationPolicy.pdf>

⁸ CDC, CDC Evaluation Framework. <http://www.cdc.gov/eval/framework/index.htm>

⁹ Department of State, DOS Program Evaluation Policy. 2012, <http://www.state.gov/s/d/rm/rls/evaluation/2012/184556.htm>

across PEPFAR programs. In addition, the ESoP responds to recommendations by the GAO and the IOM, and stipulations within congressional reauthorization to expand the utility of evaluation processes and data across PEPFAR programming for greater accountability and transparency.

Further language in the *PEPFAR Stewardship and Oversight Act Of 2013*¹⁰, Paragraph 3R, requires that we collect and report on an annual basis:

“(R) A description of program evaluations completed during the reporting period, including whether all completed evaluations have been published on a publicly available Internet website and whether any completed evaluations did not adhere to the common evaluation standards of practice published under paragraph (4).”

As an extension to the basic objective of an evaluation to determine if a program works, UNAIDS describes three primary reasons for conducting evaluations¹¹:

- 1) Program improvement, using evaluation results as feedback to program implementers to make the program function more effectively and efficiently.
- 2) Program accountability and transparency, so stakeholders and funders alike are aware of the progress of the program.
- 3) Program scale-up, disseminating results to help stakeholders and partners better understand what the program has accomplished and to replicate similar approaches in future.

While the application of the standards may vary slightly according to specific implementing agency policies and procedures, it is expected that the ESoP and the respective planning, adherence, and reporting tools will be applied across **ALL** PEPFAR funded evaluations, implementation science and operations research activities including those receiving only partial funding from PEPFAR. These steps will ensure a consistent, coordinated, and collaborative effort and will work towards the development, dissemination, and use of high-quality evaluations. Simultaneously, these standards should not be interpreted to supersede more specific requirements associated with agency or centrally managed evaluations (e.g., Public Health Evaluation, Implementation Science, Impact Evaluations).

Operating Unit (OU) representatives or other key personnel associated with evaluation and research management who identify a potential conflict between different guidance recommendations should contact the appropriate agency Point of Contact (e.g., Evaluation Point of Contact in the Program Office (USAID), ADS (CDC), program officer (NIH), contract officer, project officer) who can assist to resolve any of these issues.

¹⁰ Public Law 113–56—DEC. 2, 2013, PEPFAR Stewardship and Oversight Act of 2013 (S.1545).

¹¹ UNAIDS, 2010, Basic Terminology and Frameworks for Monitoring and Evaluation. UNAIDS Monitoring and Evaluation Fundamentals. http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/7_1-Basic-Terminology-and-Frameworks-MEF.pdf

Section I – Terms and Key Concepts

PEPFAR defines evaluation as:

*... the systematic collection and analysis of information about the activities, characteristics, and outcomes, and impacts of programs and projects.*¹²

PEPFAR has further classified the following evaluation types as indicated in Table 1. Note that in addition to the four types defined in previous versions of ESoP, in the current version, two additional types of studies are articulated: implementation science and operations research.

Table 1. Evaluation Types

(1) PROCESS EVALUATION	
<p>Definition: A type of evaluation that focuses on program/ intervention implementation, including, but not limited to access to services, whether services reach the intended population, how services are delivered, client satisfaction and perceptions about needs and services, management practices. In addition, a process evaluation might provide an understanding of cultural, socio-political, legal, and economic contexts that affect implementation of the program/intervention¹³.</p>	<p>Example Evaluation Questions:</p> <ul style="list-style-type: none"> ○ To what extent were program activities/intervention(s) implemented as intended? ○ Were programs implemented according to quality standards? ○ To what extent was the right target population reached? ○ How did socio/political/cultural factors hinder and/or facilitate program implementation? ○ To what extent did capacity building activities result in improved financial management, organizational planning, and procurement capacity, among institutions? <p>Timeframe: Assessment of the implementation process are conducted during implementation and/or immediately post-intervention.</p> <p>Data types, sources: Interviews, focus groups, audits, observations, surveys (i.e., qualitative and quantitative).</p> <p>Example Indicators: Results as compared to goals/targets, counts of clients reached, products distributed, providers trained, acceptability of services, and client satisfaction.</p>

¹² This definition is derived from Patton, M.Q., 1997. *Utilization Focused Evaluation: The New Century Text*, Sage Publication, P. 23.

¹³ UNAIDS, 2010, *Basic Terminology*, op. cit., P. 65.

(2) OUTCOME EVALUATION

<p>Definition: A type of evaluation that determines if and by how much, intervention activities or services achieved their intended outcomes¹⁴. This form of evaluation assesses the extent to which a program achieves its outcome-oriented objectives. It focuses on outputs and outcomes (including unintended effects) to judge program effectiveness but may also assess program process to understand how outcomes are produced¹⁵.</p>	<p>Example Evaluation Questions:</p> <ul style="list-style-type: none"> o To what extent did the program activities/intervention(s) reach expected short-term outcomes/changes/effects? (e.g., changes in participants' knowledge, skills, attitudes, intentions, health status) o To what extent did the program activities/intervention reach expected intermediate-term outcomes? (e.g., changes in behaviors, norms, procedures, policies, performance at the organizational and system level) o Did the program or intervention work, by what magnitude, and, if so, for whom? o Were providers who received the training more likely to effectively counsel, screen and treat patients than those who did not? o Did the program have any unintended (beneficial or adverse) effects on the target population(s)?
	<p>Timeframe: Short-term, intermediate, and/or long-term post program implementation.</p>
	<p>Data Types, Sources: Interviews, observations, focus groups, surveys, questionnaires, clinical assessments, lab values, dose effects, etc.</p>
	<p>Indicators: Knowledge, risk behavior, infection status, or other endpoints, where groups are intended to serve as representative samples of the population or sub-population.</p>

(3) ECONOMIC EVALUATION

<p>Definition: Use of applied analytical techniques to identify, measure, value and compare the costs and outcomes of alternative interventions. Economic evaluation is a systematic and transparent framework for assessing efficiency focusing on the economic costs and outcomes of alternative programs or interventions. This framework is based on a comparative analysis of both the costs (resources consumed) and outcomes (health, clinical, economic) of programs or interventions. Economic evaluations are also used to ensure that resources are used efficiently¹⁶.</p>	<p>Example Evaluation Questions:</p> <ul style="list-style-type: none"> o What are the costs and benefits of differentiated models of care? o What is the cost effectiveness of a new treatment protocol? o How can programs minimize service delivery costs? o What is the cost-effectiveness of this intervention in improving patient outcomes as compared to other treatment models?
--	---

¹⁴ UNAIDS, 2008, Basic Terminology and Frameworks for Monitoring and Evaluation. UNAIDS Monitoring and Evaluation Fundamentals

¹⁵ GAO, 2011, Performance Measurement And Evaluation Definitions and Relationships

¹⁶Chisholm, D. and D.B. Evans (2007). Economic evaluation in health: saving money or improving care? *Journal of Medical Economics*, **10**(3): p. 325-337.

	Timeframe: Economic evaluations are often conducted retrospectively, once a program has proven to be effective but before scale up. Sometimes, they can be done prospectively to ensure efficient allocation of resources.
	Data types, Sources: survey data, quantitative and qualitative program data on service delivery, cost and expenditure data
	Indicators: Program costs, participant costs, Quality adjusted life years (QALY), disability adjusted life year (DALY), Cost per infection averted, etc.

(4) IMPACT EVALUATION

Definition: Measure the change in an outcome that is attributable to a defined intervention by comparing actual impact to what would have happened in the absence of the intervention (the counterfactual scenario). Impact evaluations (IEs) are based on models of cause and effect and require a rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. There are a range of accepted approaches to applying a counterfactual analysis, though IEs in which comparisons are made between beneficiaries that are randomly assigned to either an intervention or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured to demonstrate impact ¹⁷ .	Example Evaluation Questions: <ul style="list-style-type: none"> o Does an intervention result in behavior change as a direct result of the intervention? o Does the program improve HIV or its proxy endpoint at the population-level, when properly scaled? o What are the net effects of the program in achieving long-term outcomes? (e.g., changes in prevalence, incidence, mortality, sustainability)
	Timeframe: These are generally not measured accurately in the days or months post-intervention, but often long-term post implementation.
	Data Types, Sources: Quantitative biological and/or behavioral survey data that may or may not include recipients who directly received the intervention, Epidemiological surveillance, qualitative data, etc.
	Indicators: Incidence, prevalence, mortality, transmission rates, etc.

¹⁷ PEPFAR 2014 Country Operational Guidance and PEPFAR 2012 Supplemental Guidance on Implementation Science/Impact Evaluation.

(5) IMPLEMENTATION SCIENCE

<p>Definition: Implementation research is the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and to improve the quality and effectiveness of health services, in part through the study of influences on healthcare professional and organizational behavior¹⁸.</p>	<p>Sample Research Questions:</p> <ul style="list-style-type: none"> ○ How can we best transfer scientific advances with proven efficacy and effectiveness to routine public health practice in places beyond where the approach was studied?
	<p>Timeframe: Effects assessment depends on the study question and endpoint. These may be short-term and in instances of HIV endpoints, it may be intermediate or long-term (months or years).</p>
	<p>Data types, sources: Behavioral surveys, interviews, within/between groups comparisons, audits, simulated data; mathematical modeling on cost, impact, pace and scale to guide policy and program delivery</p>
	<p>Indicators: Changes in clinical practice or services, patient experiences, or clinical/service context, etc.</p>

(6) OPERATIONS RESEARCH

<p>Definition: OR is a scientific approach to decision-making about how to design, operate, and improve programs and systems, usually under conditions requiring the allocation of scarce or finite resources. Further, it seeks to identify solutions to problems that limit program quality, efficiency and effectiveness, or to determine which alternative service delivery strategy would yield the best outcomes¹⁹.</p>	<p>Sample Research Questions:</p> <ul style="list-style-type: none"> ○ How can we improve the efficiency or effectiveness, and minimize barriers to the success, of a given program or service? ○ How can finite resources be optimized for maximum benefit of the program?
	<p>Timeframe: Effects are assessed during implementation and scale up of the intervention to monitor challenges and changes in the intervention.</p>
	<p>Data types, sources: Qualitative and quantitative data generated from brainstorming-based SWOT (strengths, weaknesses, opportunities and threats) analysis, behavioral simulation exercises, scenario analysis, system dynamics modeling and various combinations of simulation and mathematical modeling.</p>
	<p>Indicators: Evidence of program improvement, broadly defined.</p>

¹⁸ Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, 1(1), 1.

¹⁹ World Health Organization (2008). *Framework for operations and implementation research in health and disease control programs*. http://www.who.int/hiv/pub/operational/or_framework.pdf

Section II – PEPFAR Evaluation Standards of Practice

In the broader context of evaluation within PEPFAR, S/GAC expects all PEPFAR implementing agencies and those who procure and implement evaluations to commit themselves at a minimum to evaluation practices based on the standards of practice below. Though many of these practices are already incorporated into agency policies and frameworks, recent reports have determined that they are not consistently implemented²⁰. By sharing a set of common standards of practice, greater consistency and quality among PEPFAR evaluations will be achieved with the intent that stakeholders will have the confidence to utilize results for program enhancement.

The standards of practice are introduced below in the order they are likely to be applied when conducting an evaluation. For example, evaluation usually starts by engaging stakeholders, but not all stakeholders may be familiar with complex methods necessary for some evaluation designs. Such a situation requires the evaluation team to build stakeholder capacity to consider evaluation methods and effectively contribute to decisions. For other evaluation designs, such knowledge transfer may not be needed.

THE STANDARDS OF PRACTICE

1. **ENGAGE STAKEHOLDERS**
2. **CLEARLY STATE EVALUATION QUESTIONS, PURPOSE, AND OBJECTIVES**
3. **USE APPROPRIATE EVALUATION DESIGN, METHODS, AND ANALYTICAL TECHNIQUES**
4. **ADDRESS ETHICAL CONSIDERATIONS AND ASSURANCES**
5. **IDENTIFY RESOURCES AND ARTICULATE BUDGET**
6. **CONSTRUCT DATA COLLECTION AND MANAGEMENT PLANS**
7. **ENSURE APPROPRIATE EVALUATOR QUALIFICATIONS AND EVALUATION INDEPENDENCE**
8. **MONITOR THE PLANNING AND IMPLEMENTATION OF AN EVALUATION**
9. **PRODUCE QUALITY EVALUATION REPORTS**
10. **DISSEMINATE RESULTS**
11. **USE FINDINGS FOR PROGRAM IMPROVEMENT**

1. ENGAGE STAKEHOLDERS

Engage stakeholders from the beginning and throughout the evaluation to ensure the success of the evaluation and implementation of the recommendations.

Evaluation results are only relevant if they respond to the needs of a diverse range of stakeholders. Stakeholders are any persons, organizations, or institutions that have an investment in what will result from an evaluation and what will be done with the results.

There are three critical groups of stakeholders who must be considered in planning an evaluation. These include:

- 1) those persons involved in program operations (e.g., sponsors, collaborators, partners, funding officials, administrators, managers, and staff);

²⁰ GAO, 2012, op. cit.

- 2) those persons served or affected by the program (e.g., clients, family members, community organizations, academic institutions, elected officials, advocacy groups, professional associations, and staff of related or competing organizations); and
- 3) those persons who make decisions regarding the program (e.g., country leadership, policy makers, program managers, sponsors²¹)

Stakeholders should be identified and engaged in the planning stages of evaluation, including prioritizing what to evaluate, budgeting and funding decisions, identification of the evaluation questions, and dissemination and use of findings and recommendations²². For evaluations of PEPFAR-funded programs, it is essential to involve governmental and non-governmental stakeholders from the country in which the evaluation is conducted.

The scope and level of stakeholder involvement will vary for each program evaluation. For example, some stakeholders, such as evaluators in the community, can be directly involved in designing and conducting the evaluation. Other stakeholders, such as policy makers, may be involved in the initial framing of evaluation questions, and they should be kept informed of the progress of the evaluation and of the evaluation results. Configuring a communication strategy at the start of program planning is particularly relevant for evaluation efforts to ensure ongoing stakeholder engagement and support, particularly if an evaluation crosses organizational units with overlapping or complementary missions.

2. CLEARLY STATE EVALUATION QUESTIONS, PURPOSE, AND OBJECTIVES

Make explicit the evaluation questions, purpose, and objectives. Evaluation planning should be part of program planning from the start and throughout program implementation.

The first stage of planning an evaluation is the general purpose of the evaluation, the specific objectives, and the specific questions. Even though methods and analyses may be complex, the purpose and objectives must be understood from the beginning, and the evaluation questions should be simple and clear. The next step should be to synthesize the best available evidence regarding the intervention(s); i.e. what is already known about how well it works. Focus should then shift to:

- 1) what will be evaluated (specifically defining the intervention or aspects of an intervention),
- 2) who wants the information,
- 3) what do they want to know (various outputs or outcomes), and
- 4) how the results will be linked to specific future decisions or programs²³

Clarifying the intent and answers to these questions from the beginning will facilitate the subsequent decisions. One should anticipate that the planning stage of an evaluation would require several iterations and extensive review, ultimately serving to refine the evaluation to ensure that the evaluation questions, methods, and analyses are appropriate and sound. Following these developments, other steps in evaluation implementation will move quickly.

²¹ Centers for Disease Control and Prevention. Framework for program evaluation in public health. MMWR 1999;48 (No. RR-11)

²² Measure Evaluation, 2011. Tools for Data Demand and Use in the Health Sector: Stakeholder Engagement Tool (<http://www.cpc.unc.edu/measure/publications/ms-11-46-e>).

²³ USAID, Automated Directives System 203.3.1.4 11_02_2012

3. USE APPROPRIATE EVALUATION DESIGN, METHODS, AND ANALYTICAL TECHNIQUES

Knowing the program maturity, the questions to be addressed, and the resources available are requisite to determine the appropriate evaluation design, methods and analytical techniques.

The evaluation design and methods should be feasible, context sensitive, culturally relevant, and rigorous. When selecting the design and methods, consider the program's maturity, the evaluation questions to be addressed, purpose and timeline for needing the results, the intended audience, and the available financial and other resources. As noted below in Standard of Practice 5, the design-resource correspondence is critical, since a pre-determined budget may not support an evaluation design sufficient to address important questions.

In conjunction with evaluation design and methods, an analysis plan should be pre-determined and described in an evaluation Scope of Work (SOW)/protocol. Because evaluations generally address multiple questions, a range of analytic methods is often needed. For example, in many instances a mixed-methods approach that combines quantitative and qualitative methods and analyses is ideal. One needs to use the most rigorous methods appropriate for the evaluation questions. Carefully thinking through data needs and analytic techniques in advance will enhance the quality, credibility, and usefulness of an evaluation by increasing the strength and specificity of the findings and recommendations.

4. ADDRESS ETHICAL CONSIDERATIONS AND ASSURANCES

Address human rights protections when planning and implementing the evaluation.

Evaluations must be conducted in a manner that is respectful to and protects human rights, privacy, and confidentiality, and maintains the dignity of participants and other stakeholders²⁴. U.S. government (USG) agencies follow regulatory standards regarding human protections²⁵ and these standards are based on principles and guidelines established in the international community²⁶. These principles require that evaluators behave legally, ethically, and have regard for the physical and psychological welfare of those involved and those affected by an evaluation, including vulnerable populations. All personnel involved in planning and implementing an evaluation should be knowledgeable regarding agency policies, rules and regulations in this regard, and complete ethical certifications when indicated.

Evaluation procedures should ensure that participants who contribute information to a study, especially if they may incur risks, do so willingly and with full knowledge of any potential risks. Participants should be informed that their eligibility to receive services is not contingent upon their participation in the evaluation (i.e., clients retain the right to refuse to participate in an evaluation). These protections are generally addressed in an informed consent agreement administered before participants agree to participate and respond to data collection inquiries. Such protections also should be described in the evaluation protocol governing the conduct of the evaluation.

Special protections are especially important when conducting evaluations involving children, prisoners,

²⁴ AEA, 2010, Roadmap, *op.cit.*

²⁵ Department of Health and Human Services (45 CFR 46), United States Agency for International Development (22 CFR 225).

²⁶ The WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects (<http://www.wma.net/en/30publications/10policies/b3/>) and the Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines for Biomedical Research Involving Human Subjects (http://www.cioms.ch/publications/layout_guide2002.pdf) are the documents used by agencies to articulate their rules and regulations regarding Human Subjects.

pregnant women, and other vulnerable groups.

Depending on the objectives, questions, and methods of the evaluation, evaluation scopes of work/protocols may have to go through an Institutional Review Board (IRB) at the in-country national level, at a USG implementing agency, and when relevant, at the level of the implementing partner or the associated external institutional entity.

5. IDENTIFY RESOURCES AND ARTICULATE BUDGET

Identify the evaluation budget at the start of program planning.

Successful execution of evaluations requires not only a commitment among program managers and implementers to incorporate evaluation into their efforts, but also sufficient resources to ensure the implementation of the appropriate type and design of evaluation. This commitment requires considerable forethought, since resource decisions are often made in the context of tension between program and evaluation priorities. As planning occurs, these priorities need to be balanced and adequate resources need to be made available.

A recommended range of resource commitments for monitoring and evaluation are available in the literature and guidance materials, but it is important to note that these ranges typically apply to **monitoring** activities only²⁷. Funding for formal **evaluation** efforts, including process, outcome, impact, or economic evaluations, may require additional allocations above these ranges. Evaluation planning early in the program planning process should provide a reasonable estimate of these funding requirements, and appropriate steps need to be taken to ensure resources are available to fulfill the requirements of the evaluation design.

6. CONSTRUCT DATA COLLECTION AND MANAGEMENT PLANS

Create data collection and management plans prior to implementing the evaluation to ensure that data are valid, reliable, and accessible.

Since the intent of data collection is to gather information that stakeholders perceive as trustworthy and relevant, evaluation scopes of work/protocols should include a data collection and management procedure that is in line with agency policies and specifies the following: who will administer the data collection instruments; when these will be administered; how data will be gathered and checked in a systematic, comparable, precise, and unbiased way so that data are accurate, complete, and valid; how the data will be archived, transported, secured, confidentiality ensured, and disposed of (if applicable); how data-use agreements will be developed with partners and others; which institutions and individuals will have access to the data in its various forms; how long the data will be saved; how interview tapes or audio files will be managed and stored; whether or not they will be translated, transcribed; and how observations will be captured and stored.²⁸ The plan might include an agreement signed by evaluation team members that acknowledges their responsibilities in this area.

²⁷ In PEPFAR COP Guidance, the recommended range applies to the M&E of program implementation, and the same range applies to the larger country COP budget to support system strengthening work for SI. In the context of program implementation, agencies will make the determination of how to apply evaluation requirements and funding. For example, evaluations may be conducted at project sites, for a program, or for an entire implementing mechanism.

²⁸ Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users* (3rd ed.). Thousand Oaks, CA: Sage.

7. ENSURE APPROPRIATE EVALUATOR QUALIFICATIONS AND INDEPENDENCE

Ensure that an evaluator has appropriate experience and capabilities. Manage any conflicts of interest of the evaluators (or team) and mitigate any untoward pressures that could be applied to the evaluator or evaluation team that would influence its independence.

It is important that the evaluation team members:

- are qualified to conduct the evaluation through knowledge and experience;
- disclose any potential conflict of interest with the evaluation;
- are protected from any undue pressure or influence that would affect the independence of the evaluation or objectivity of the evaluator(s).

Only evaluation teams (whether internal or external) that possess the education, capabilities, skills and experience appropriate to undertake the tasks proposed in the SOW/protocol should conduct evaluations. Professional evaluators typically have advanced training in social science or public health, and depending on the nature of the program and the evaluation questions, the evaluation team might also require members with specialized subject area expertise (e.g., epidemiology, clinical skills, economics, statistics, qualitative analysis).

The résumés of the evaluation team members should be examined both to assess appropriateness of their skills/competencies, regional expertise, languages spoken, subject matter expertise (HIV Testing and Counseling, Key populations, male circumcision), as well as for the appropriate education and previous experience(s) to perform the specific evaluation at hand.

It is vital to manage any conflicts of interest of the evaluator and the evaluation team to ensure credibility and mitigate bias. In advance, everyone on the evaluation team must disclose any personal, financial, or other relationships they have that might pose a conflict of interest (or the appearance of a conflict) in their role as evaluators. This is frequently accomplished by having the evaluation team sign a conflict of interest (COI) statement prior to conducting the evaluation. The COI statements should be kept with all other evaluation data and shown to stakeholders as appropriate. The COI statements should be in line with the implementing agency conflict of interest policy (if there is one) and should be included in both the protocol/SOW and the appendices of the final evaluation report.

Managing the independence of the evaluation includes informing and educating all those participating in the evaluation (including those collecting data, funding, reviewing, or approving the evaluation) that the planning, implementation and results of the evaluation should not be manipulated in any way to suggest undue influence. Suggested strategies to improve evaluator independence include, but are not limited to, having evaluation units that are separate from program units; using external evaluations and evaluators; or establishing formal conflict-of-interest procedures and declarations for internal and external evaluators. In some instances, if certain procedures or activities are likely to produce misleading information or conclusions, the evaluation team has the responsibility to communicate their concerns to relevant stakeholders and colleagues. The team should identify proper ways to proceed (e.g., discussions at a higher level, a dissenting cover letter or appendix, refusal to sign the final report, documenting concern and make a disclaimer, or submitting a Statement of Difference letter^{29 30 31}).

²⁹ USAID-specific procedure

³⁰ World Bank, Independence and Impartiality in Conducting Evaluations, Chap 3 in XXXX. World Bank, Washington, D.C. http://siteresources.worldbank.org/EXTGLOREGPARPROG/Resources/grpp_sourcebook_chap3.pdf

³¹ American Evaluation Association (AEA), Guiding Principles for Evaluators, <http://www.eval.org/p/cm/ld/fid=51>

8. MONITOR THE PLANNING AND IMPLEMENTATION OF AN EVALUATION

Continuous, planned monitoring of the evaluation is important to the successful completion of the evaluation.

Monitoring how an evaluation is planned and implemented is essential for ensuring quality evaluations; controlling redundancy, time and costs; and identifying and solving unexpected problems as they arise. This monitoring activity should start with the appropriate agency lead, the evaluation lead and relevant stakeholders tracking final development and completion of the evaluation protocol. Once the evaluation has begun, it is important to document progress in accordance with the evaluation design, and especially any contextual changes, deviations from the evaluation plan, or quality on a regular basis. All relevant stakeholders should be kept informed of the evaluation progress.

Data limitations or new information about the project or program being evaluated may arise as the evaluation is being conducted and this may have implications on the adequacy of the original plan or the feasibility of answering the evaluation questions. If adjustments are necessary, the evaluation team should document these changes, along with the rationale, and submit modifications for approval depending on relevant IRB and USG agency requirements. Any modifications should be reflected in midterm and final reports. Monitoring and documenting the progress of an evaluation and communicating with stakeholders is a primary responsibility of the evaluation lead(s). The evaluation team simultaneously has the responsibility for safeguarding its quality, adhering to the SOW/protocol, and applying the evaluation standards throughout.

9. PRODUCE QUALITY EVALUATION REPORTS

The final evaluation report should contain certain elements to ensure the quality and transparency of the evaluation.

The evaluation report should represent a thoughtful, methodologically sound, and well-organized effort to describe the process and findings of the evaluation of a PEPFAR program. The content of an evaluation report should allow the reader to fully understand the context/background, the rationale for conducting the evaluation, the evaluation questions to be answered, and the methods and analyses that were used. The report should assess the findings of the project/program, supported by strong quantitative, qualitative, or mixed methods evidence. Any limitations to the evaluation findings (e.g., spillover, lost to follow up, or poor quality or missing data) should be fully described. For the findings to be useful, supported recommendations should be made explicit and be actionable, realistic, and specific.

Developing and using templates for evaluation reporting can serve to standardize evaluation reports and ensure that all required components of an evaluation report are included. Appendix C of this ESoP document provides a checklist of a minimum set of these elements for all PEPFAR-funded evaluations. This checklist also can be used as a basis for peer review of the final report, to help improve report quality. Critically, final reports provide the information necessary for the formal review of evaluations for compliance with the standards of practice; consequently, inclusion of all of the required elements is essential³².

³² The **PEPFAR Evaluation Standards of Practice, Version 1.0**, 2014, made an allowance for published articles to function as final reports. This accommodation is no longer supported.

10. DISSEMINATE RESULTS

Evaluation results should be disseminated to all stakeholders, the public and funders.

Evaluation results and recommendations should be presented clearly and simply so that stakeholders and other parties can easily understand the evaluation process, results and recommendations³³. Like other elements of evaluation, dissemination should be discussed and planned in advance with stakeholders and must follow agency evaluation dissemination instructions. These steps will ensure that the information needs of relevant audiences will be met, which requires consideration of the timing, style, tone, message source, vehicle and format of information products (e.g., publications, briefings, newsletter).

Evaluation results can be shared via evaluation reports, publications, oral presentations, agency websites, annual reports, and briefings. As congressionally mandated in the *PEPFAR Stewardship and Oversight Act*,³⁴ all completed evaluation reports must be published on a publically available internet website. All completed PEPFAR evaluations must be uploaded onto the implementing agency databases within 90 days after approvals of all relevant authorities. These agency databases will constitute the core of this dissemination approach, and additional access will be provided through linkages from a central PEPFAR site (e.g., www.PEPFAR.gov).

11. USE FINDINGS FOR PROGRAM IMPROVEMENT

Evaluation findings and recommendations should be utilized for decision-making and program improvement.

Evaluation is a worthwhile endeavor only if the results are used. Well-planned evaluations provide evidence to inform real-world decision-making and contribute to learning agendas that have national, regional, or global importance. Evaluation results can be used to demonstrate the effectiveness of a program, plan new procurements, make mid-stream adjustments to improve programs, and demonstrate accountability. The evaluation objective(s) as well as the intended audience identified in the planning phase of the evaluation should guide use of the evaluation results. These various agendas for use of evaluation findings also should be stipulated in the evaluation plan.

Evaluation findings can be more useful if:

1. the evaluation question was linked to a specific future programmatic decision;
2. intended users are engaged early, their information needs are identified, and recommendations are made according to the user needs;
3. intended users review evaluation reporting and dissemination plans;
4. intended users are reminded of the planned use of the evaluation findings;
5. evaluation findings and recommendations are translated into usable information products targeted to specific intended users;
6. intended users and stakeholders are supported in applying the findings and recommendations;
7. intended users are supported in making strategic choices about where to focus follow-up efforts;
8. implementation of the recommendations is monitored; and intended users and other stakeholders see evaluation as an ongoing process rather than a one-time event or moment-in-time report³⁵.

³³ GAO, 2012, *Designing Evaluations, Revision*, GAO-12-208G. Washington, D.C. January 2012.

³⁴ Public Law 113-56, *op. cit.*

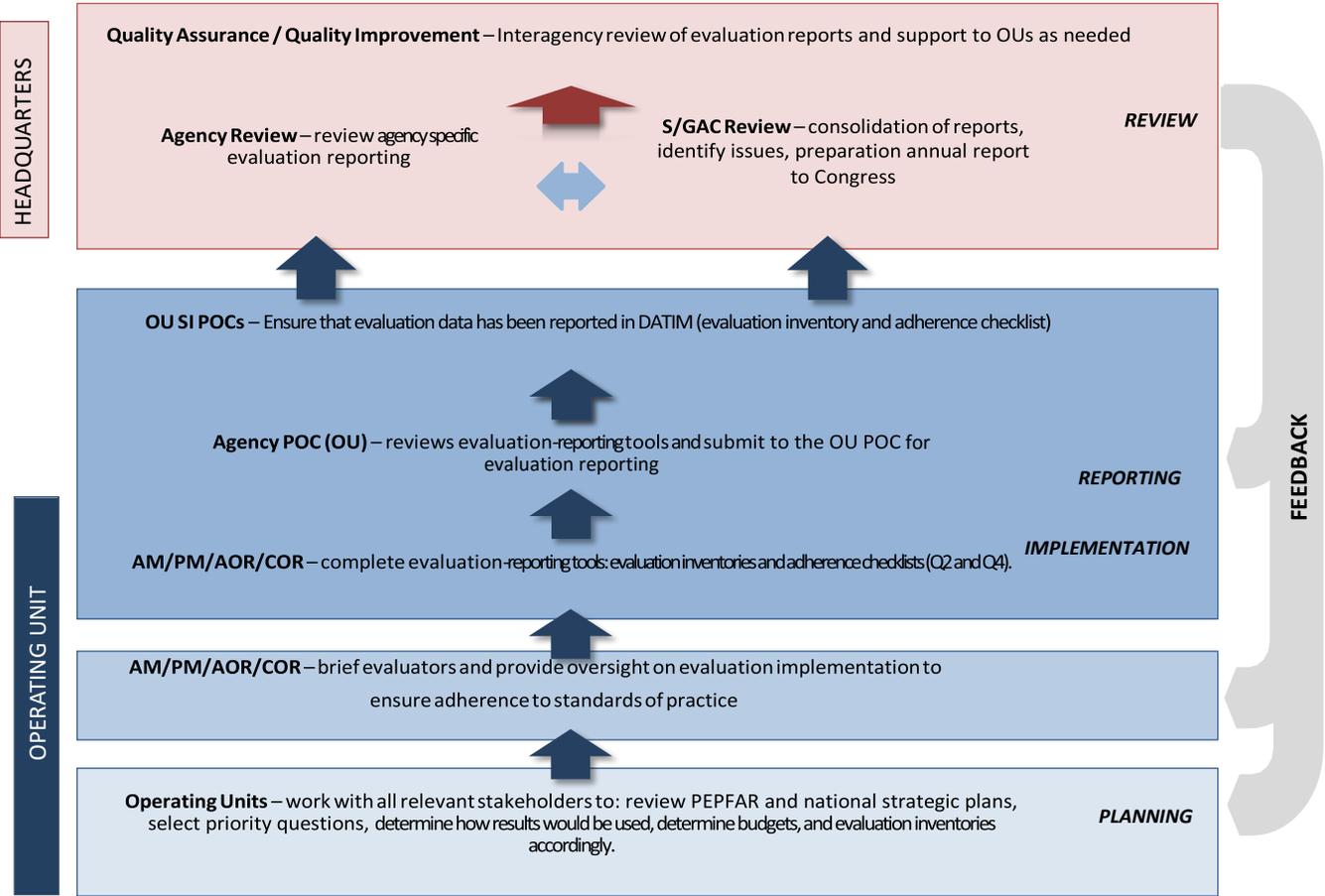
³⁵ Patton, M.Q., 2008. *Utilization-focused evaluation*, 4th edition. Thousand Oaks, CA: Sage.

Section III - Planning, Implementation, Reporting, and Review

As USG country teams approach strategic planning for evaluation portfolios (including implementation science and operations research), it is critical to keep in mind that **ALL** PEPFAR-funded evaluations and operations research/implementation science activities— whether conducted by implementing partners, evaluation partners, agency personnel, and whether using headquarter (headquarter operating plan), central funds (e.g., PHE, Implementation Science) or country funds (country operating plan or other bilateral funds), all of them need to be situated within the larger context.

Evaluation and research priorities defined at the country level based on information needs should guide this approach. Interagency collaboration, as well as completeness and transparency of information are essential to inform this strategic planning, and in addition select information regarding all PEPFAR funded evaluations is required to report to headquarters and Congress. This section describes the entire cycle of evaluation planning, implementation, reporting, and review (Figure 1).

Figure 1. Evaluation planning, implementation, reporting, and review process



EVALUATION PLANNING

Evaluation planning is important to ensure that evaluation resources are allocated appropriately and in concert with overall evaluation priorities³⁶. The American Evaluation Association targeting US Federal Agencies states the following regarding evaluation planning, from *An Evaluation Roadmap for a More Effective Government*.

...Each Federal agency should require its major program components to prepare annual and multi-year evaluation plans of the studies and other evaluation activities that they will undertake. The plans should be updated annually. The planning should take into account the needs of evaluation results for informing program budgeting, reauthorization, agency strategic plans, ongoing program development and management and responses to critical issues that arise concerning program effectiveness, efficiency, and waste. These plans should include an appropriate mix of short and long term studies to ensure that evaluation results of appropriate scope and rigor are available when short or long term policy or management decisions must be made. To the extent practical, the plans should be developed in consultation with program stakeholders who are involved in or affected by the programs³⁷.

DEVELOPING THE EVALUATION INVENTORY

While each Agency and OU has flexibility in how it develops its evaluation inventory, the following steps are suggested.

- a. Review existing evaluation inventory and ensure that all planned and on-going evaluations remain relevant and if not, determine if they are still needed.
- b. Determine how evaluation results will be used and by whom. For every evaluation, there should be a clearly articulated audience and purpose/use of results.
- c. Evaluation questions should guide the evaluation design, selection of data sources, methods, data collection and analysis, budget, and utilization of evaluation results for program improvement.
- d. Ensure that an adequate budget is allocated towards the evaluation. Items to be taken into considerations for budgeting are: types of expertise required, whether the evaluation will be contracted externally and if so, the cost implications, geographic scope and purpose, and sampling to name a few. Recommended percentages in the literature for evaluation vary from 3% to 10% of program budget.

It is recognized that as much strategic planning as it done, evaluation questions can emerge unexpectedly and urgently. This therefore requires a certain degree of flexibility to ensure that emerging questions and information needs can be addressed accordingly.

EVALUATION IMPLEMENTATION

Evaluators should implement the evaluation as per the SOW/protocol in close communication with the AM/PM/AOR/COR/IAD that commissioned the evaluation. The AM/PM/AOR/COR/IAD will provide evaluators with the ESoP to ensure that they are aware of the standards and associated measures of adherence. The

³⁶ GAO, 2012, PEPFAR Evaluation, *op. cit.*

³⁷ American Evaluation Association, 2010, Roadmap, *op. cit.*, P. 6.

AM/PM/AOR/COR/IAD should provide oversight of the evaluation implementation including in-briefing, reviewing work plans and data collection tools, ensuring ethical procedures, participating in the site selection, and de-briefing with the evaluation team. Upon completion of evaluations, evaluators will use, at a minimum, the Required Components of an Evaluation Report (Appendix C) to develop a final evaluation report. The reports must be disseminated publically within 90 days after agency approval of the final evaluation report.

EVALUATION REPORTING

In order to report on evaluations, OUs must complete the Evaluation Inventory and the Adherence Checklist on all completed evaluations. A data dictionary is available on pepfar.net with details and specific instructions on reporting.

Evaluation Inventory The inventory is where **all** newly commencing, ongoing, and completed evaluations, implementation science and operations research activities are reported for the current fiscal year. Your responses to the evaluation inventory will be used in the Annual Report to Congress.

Adherence Checklist The checklist must be completed for all **completed** evaluations and OR/IS activities within 90 days after approval of the final report by the Agency. The primary purpose of the checklist is to assess adherence of PEPFAR evaluations to the 11 standards. The final report should be used as the primary reference document to complete this section of the tool.

EVALUATION REVIEW

Review processes may vary by Agency but should include reviews at different levels at the Agency, OU, and S/GAC. Once evaluation reporting has been completed and submitted to S/GAC via DATIM, the appropriate headquarter agency POCs will examine data for completeness. For the data for which there is incomplete information, the Agency Evaluation POCs will work with missions to ensure completeness and accuracy.

Common errors from past years of evaluation reporting include:

- **Missing evaluation title:** Enter evaluation title (enter a tentative title if none exists yet).
- **Incorrect evaluation type:** Review definitions of each evaluation type and seek support from agency evaluation POC as needed.
- **Missing/poorly-worded evaluation questions:** Enter evaluation questions into evaluation questions field; seek support from agency evaluation POC as needed.
- **Evaluations without dates:** There are now two date fields, for start and end dates including (proposed and actual). Enter both start and end dates; update end date as needed.
- **Evaluations lasting longer than 5 years:** Update fields annually, including evaluation stage, evaluation progress, and cost of project to date.
- **Evaluation listed as “ongoing” with passed end date:** If evaluation has ended, ensure all EI fields are updated and an AC checklist is completed. If evaluation has not ended, update end date field in inventory to anticipated end date.

- **Evaluation budget missing/inconsistently reported:** Enter evaluation budget information correctly. Seek information from budget staff as needed.
- **General:** Lack of updating fields annually (e.g., evaluation stage, evaluation progress).
- **General:** Evaluations not entered into Evaluation Inventory. Need to enter every PEPFAR evaluation, seek support from agency evaluation POC as needed.

Common errors regarding the checklist are that one is not filled for every completed evaluation. S/GAC will compile the evaluation data facilitate an interagency review. Data drawn from the Adherence Checklists will be incorporated into the Annual Report to Congress, complying with the legislative requirements of the PEPFAR Stewardship and Oversight Act. We hope that improvements in the Evaluation Inventory and in the related supplemental materials such as the data dictionary will help to clarify and reduce errors in evaluation reporting.

Section IV- Roles and Responsibilities

For evaluation to become integrated into PEPFAR programs, S/GAC, Implementing Agencies, and OUs have specific responsibilities, as well as extensive shared interests. Among all of these partners, work needs to be coordinated to preclude unnecessary duplication. Communication channels should be established among the various Points of Contact in the field and at HQ to support coordination and improve the quality of evaluation implementation.

It is essential to recognize that Implementing Agencies have existing business practices, evaluation guidance, policies and frameworks^{38,39} which are generally inclusive of the standards of practice described. These variations should be articulated in agency-specific translational documents providing guidance on the integration of ESoP requirements with agency-specific policies and procedures. Table 2 illustrates ESoP roles and responsibilities by stakeholder entity, and Table 3 illustrates roles and responsibilities for individuals.

³⁸ USAID, Evaluation Policy, *op. cit.*

³⁹ CDC, Evaluation Framework, *op. cit.*

Table 2. Stakeholder roles and responsibilities

Stakeholders	Roles & Responsibilities
S/GAC	<ul style="list-style-type: none"> • Coordinate interagency process to support implementation of evaluation standards of practice and use of high quality evaluations in PEPFAR-supported programs. • Coordinate interagency evaluation experts to develop and disseminate guidance, orient the field on the ESoP, and respond to high-level inquiries from the field and headquarters. • Consolidate and review agency and OU reporting of all evaluations and ESoP adherence checklists. • Coordinate efforts to provide technical assistance to strengthen agency, OU, and implementing partner evaluation capacity, as appropriate. • Coordinate interagency efforts to disseminate lessons learned, best practices, and evaluation results to relevant stakeholders including the public to promote use of findings and improve PEPFAR programs. • Coordinate public dissemination of evaluation reports via pepfar.gov. • Coordinate the development of an annual report to Congress on the descriptions of all PEPFAR funded evaluations, public dissemination of evaluation reports, and adherence to ESoP.
Interagency Evaluation Experts (short-term task team)	<ul style="list-style-type: none"> • Develop and update ESoP and orient the field and headquarter staff through webinars and other dissemination fora. • Serve as a resource group to answer questions and provide technical assistance to OUs and headquarters regarding implementation and use of the ESoP, as appropriate. • Review and revise ESoP and relevant tools and templates as needed. • Collaborate and support S/GAC activities, as needed. • Orient the field on ESoP reporting requirements in APR as well as COP/HOP and other planning requirements.

Implementing Agencies	<ul style="list-style-type: none"> • Identify Implementing Agency Evaluation Point of Contact at headquarters and in each OU who will be responsible for planning, organizing technical assistance, collecting relevant data for all evaluations (inclusive of centrally funded, multi-country, and other evaluations), and reporting on ESoP. • Support dissemination and orientation of ESoP to OUs. • Develop agency-specific guidance documents to translate PEPFAR ESoP for integration with agency policies, when appropriate. • Lead agency efforts to ensure and improve adherence to the ESoP through orientation/training, rollout, oversight, as well as protocol / SOW and evaluation report review. • Contribute to completion and review of Evaluation inventories and Adherence Checklists in OUs, and support submission for APR reporting. • Provide agency-specific technical assistance to OUs to strengthen evaluation capacity, and participate in interagency technical assistance when appropriate. • Disseminate lessons learned, best practices, and evaluation results to relevant stakeholders and the public. • Disseminate all completed PEPFAR evaluation reports (e.g., OU-implemented, agency- implemented, headquarter-implemented) on publically accessible agency website and submit the web addresses to S/GAC. • Support use of evaluation results for decision-making and program improvement within the agency.
Stakeholders	Roles & Responsibilities
Operating Units	<ul style="list-style-type: none"> • Understand and utilize the ESoP for evaluation planning, implementation, and reporting. • Ensure that all evaluations are entered into the Evaluation Inventory form. • Ensure that all completed PEPFAR evaluations and IS/OR activities (e.g., implementing partner, agency, central, multi-country) are described and assessed for adherence to the standards using the Adherence Checklist. • Ensure that completed Adherence Checklists are completed within expected timeframes (i.e., 90 days after agency approval of final evaluation report). • Ensure that all planned PEPFAR evaluations are designed and budgeted appropriately during Country Operational Planning process. • Facilitate capacity building needs of staff within the OU on evaluation planning, implementation, budgeting, and oversight. • Request ESoP-related technical assistance through Agency POC who will coordinate with interagency evaluation experts and/or others to provide support as needed. • Ensure that all OU final evaluation reports are disseminated on relevant agency websites within 90 days of completion and submit web addresses to S/GAC. • Disseminate lessons learned, best practices, and evaluation results with relevant in-country stakeholders and the public as appropriate. • Ensure use of evaluation results for decision-making and program improvement.

Evaluators	<ul style="list-style-type: none">• Collaborate with relevant implementing agency and external or internal agency evaluators to provide access to data, information, human resources, and sites for evaluation efforts.• Apply ESoP when developing and implementing PEPFAR-funded evaluations.• Participate, if applicable, in evaluation process as key informants regarding the projects and interventions being implemented.• Use evaluation findings for program improvement.
-------------------	---

Table 3. Individual roles and responsibilities

Individuals	Roles & Responsibilities
Implementing Agency ESoP POC (HQ)	<ul style="list-style-type: none"> • Support dissemination and orientation of ESoP to OUs. • Develop agency-specific guidance documents to translate ESoP for integration with agency policies, when appropriate. • Provide agency-specific technical assistance to OUs to strengthen evaluation planning, implementation, and reporting. • Ensure dissemination of completed PEPFAR evaluation reports on agency website and submission of agency website link to S/GAC.
Implementing Agency ESoP POC (OU)	<ul style="list-style-type: none"> • Facilitate strategic evaluation planning at OU level. • Collect relevant data regarding planned evaluations from the AM/PM/AOR/COR and submit • Collect relevant data on newly commencing, ongoing, and completed evaluations from AM/PM/AOR/COR and submit to the SI Liaisons during APR. • Submit completed evaluation reports to Implementing Agency HQ ESoP POC.
AM/PM/AOR/COR or Implementing Agency Designee	<ul style="list-style-type: none"> • Identify which projects will be evaluated and provide relevant information to the SI Liaison during COP. • Develop evaluation Scopes of Work or protocols, which must be approved through relevant agency procedures and processes. • Select and procure the services of competent and qualified evaluators to conduct evaluations. • Provide management / oversight of evaluators, providing them with the ESoP including relevant tools and templates. • After the evaluation/OR IS activity is implemented, complete the Evaluation Inventory and, when appropriate, the Adherence Checklist within 90 days after Agency approval of the final evaluation report. • Submit completed inventory and checklists to SI Liaisons, through Agency POC.
Evaluator(s)	<ul style="list-style-type: none"> • Review and understand PEPFAR ESoP. • Ensure evaluations adhere to the 11 standards. • Provide completed, high quality evaluation reports ensuring the inclusion at least of the Required Components of an Evaluation Report. • Provide findings, conclusions, and recommendations, which are feasible, actionable, and specific to inform program planning and facilitate program improvement.

Appendix A: Definition of Terms⁴⁰

ACTIVITY: “An activity is a component of a project that contributes to a project purpose. It refers to an award (such as a contract, grant or cooperative agreement), or a component of a project such as training or technical assistance.”⁴¹

CONFLICT OF INTEREST: “A situation in which a party has interests that could improperly influence that party’s performance of official duties or responsibilities, contractual obligations, or compliance with applicable laws and regulations”⁴². A real or perceived conflict of interest of an evaluator translates to a lack of “impartiality, objectivity, and integrity”⁴³ and could jeopardize the credibility and validity of the findings.

ECONOMIC EVALUATION: Use of applied analytical techniques to identify, measure, value and compare the costs and outcomes of alternative interventions. Economic evaluation is a systematic and transparent framework for assessing efficiency focusing on the economic costs and outcomes of alternative programs or interventions. This framework is based on a comparative analysis of both the costs (resources consumed) and outcomes (health, clinical, economic) of programs or interventions. Main types of economic evaluation are cost minimization analysis (CMA), cost-effectiveness analysis (CEA), cost-benefit analysis (CBA) and cost-utility analysis (CUA)⁴⁴.

EVALUATION: “Evaluation is the systematic collection and analysis of information about the characteristics and outcomes of programs and projects as a basis for judgments, to improve effectiveness, and/or inform decisions about current and future programming. Evaluation is distinct from assessment, which may be designed to examine country or sector context to inform project design, or an informal review of projects.”⁴⁵

EVALUATOR INDEPENDENCE: There are multiple facets pertaining to independence that need to be considered. One facet refers to having “no fiduciary relationship with the implementing partner” that is being evaluated”⁴⁶. In addition, “independence provides legitimacy to evaluation and reduces the potential for conflict of interest which could arise if policy makers and managers were solely responsible for evaluating their own activities”⁴⁷. Also “evaluators are independent from the development intervention, including its policy, operations and management functions, as well as intended beneficiaries...The evaluation team is able to work freely and without interference. It is assured of co-operation and access to all relevant information”⁴⁸

EXTERNAL EVALUATION: An evaluation conducted externally by an independent consulting firm, research institute, or independent oversight agency such as GAO or an agency’s Inspector General. The importance of an evaluator’s independence from program management provides greater credibility of the evaluation findings and report.⁴⁹

⁴⁰ Agencies (as well as global partners) use generally comparable definitions for these terms, but some variation does exist and may have implications for specific work performed.

⁴¹ Department of State, Evaluation Policy, *op. cit.*

⁴² Asian Development Bank. 2005. Guidelines to avoid conflict of interest in independent evaluations. P. 46.

⁴³ *Ibid.*

⁴⁴ Drummond 2005.

⁴⁵ *Ibid.*

⁴⁶ United States Agency for International Development, ADS 203. P. 10.

⁴⁷ OECD DAC, 1991. Principles for Evaluation of Development Assistance. P. 6.

⁴⁸ OECD DAC, Quality Standards for Development Evaluation. P. 11

⁴⁹ GAO, 2012. Designing Evaluations, 2012 Revisions. P. 5

IMPACT: “The long-term, cumulative effect of programs/interventions over time on what they ultimately aim to change, such as a change in HIV infection, AIDS-related morbidity and mortality. Note: Impacts at a population-level are rarely attributable to a single program/ intervention, but a specific program/intervention may, together with other programs/interventions, contribute to impacts on a population.”⁵⁰

IMPLEMENTATION SCIENCE: Implementation research is the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and to improve the quality and effectiveness of health services, in part through the study of influences on healthcare professional and organizational behavior⁵¹.

IMPACT EVALUATION: Measures the change in an outcome that is attributable to a defined intervention by comparing actual impact to what would have happened in the absence of the intervention (the counterfactual scenario). IEs are based on models of cause and effect and require a rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. There are a range of accepted approaches to applying a counterfactual analysis, though IEs in which comparisons are made between beneficiaries that are randomly assigned to either an intervention or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured to demonstrate impact.⁵²

INTERNAL EVALUATION: Evaluations aimed at identifying program improvement and are conducted by a program office or an agency unit that specializes in program analysis and evaluation.⁵³ Internal evaluations include those led by or made up entirely of implementing agency staff (HQ or field), those implemented by partners of their own efforts, or those commissioned by implementing partners using external consultants.

MONITORING: “Monitoring provides an indication of progress against goals and indicators of performance, reveals whether desired results are occurring, and confirms whether implementation is on track. In general the results measured are the direct and near term consequences of program activities.”⁵⁴

OPERATIONS RESEARCH: Operations research is a scientific approach to decision-making about how to design, operate, and improve programs and systems, usually under conditions requiring the allocation of scarce or finite resources. Further, it seeks to identify solutions to problems that limit program quality, efficiency and effectiveness, or to determine which alternative service delivery strategy would yield the best outcome⁵⁵.

OUTCOME: “Short-term or medium-term effect of an intervention’s outputs, such as a change in knowledge, attitudes, beliefs, behaviors.”⁵⁶

OUTCOME EVALUATION: “A type of evaluation that determines if and by how much, intervention activities or services achieved their intended outcomes.” It focuses on “outputs and outcomes (including unintended

⁵⁰ UNAIDS, 2010, Basic Terminology and Frameworks for Monitoring and Evaluation, *op. cit.*, P. 62.

⁵¹ Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, 1(1), 1.

⁵² PEPFAR 2014 Country Operational Guidance, and PEPFAR 2012 Supplemental Guidance on Implementation Science/Impact Evaluation.

⁵³ GAO, 2012, *op. cit.* P. 5

⁵⁴ *Ibid.*

⁵⁵ World Health Organization (2008). Framework for operations and implementation research in health and disease control programs. http://www.who.int/hiv/pub/operational/or_framework.pdf

⁵⁶ UNAIDS, 2010, Basic Terminology, *op. cit.*, P. 65.

effects) to judge program effectiveness, but may also assess program process to understand how outcomes are produced.”⁵⁷

OUTPUTS: “The results of program/intervention activities; the direct products or deliverables of program/intervention activities, such as the number of HIV counseling sessions completed, the number of people served, the number of condoms distributed.”⁵⁸

PROCESS EVALUATION: “A type of evaluation that focuses on program or intervention implementation, including, but not limited to access to services, whether services reach the intended population, how services are delivered, client satisfaction and perceptions about needs and services, management practices. In addition, a process evaluation might provide an understanding of cultural, socio-political, legal, and economic context that affect implementation of the program or intervention.”⁵⁹

PROGRAM: “An overarching national or sub-national response to a disease. A program generally includes a set of interventions marshaled to attain specific global, regional, country, or subnational objectives; involves multiple activities that may cut across sectors, themes and/or geographic areas.”⁶⁰

PROJECT: “An intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader program.”⁶¹

PROTOCOL: “A study protocol is a document that describes, in detail, the plan for conducting the [‘clinical’ – in the original] study. The study protocol explains the purpose and function of the study as well as how to carry it out. Some specific things included in the protocol are the reason for the study, the number of participants, eligibility and exclusion criteria, details of the intervention or therapy the participants will receive (such as frequency and dosages), what data will be gathered, what demographic information about the participants will be gathered, steps for clinical caregivers to carry out, and the study endpoints. A single standard protocol must be used without deviation to ensure that the resulting data will be significant and reliable”.⁶² We refer to a protocol for an evaluation as opposed to a clinical study.

RESEARCH: “A systematic, intensive study intended to increase knowledge or understanding of the subject studied, a systematic study specifically directed toward applying new knowledge to meet a recognized need, or a systematic application of knowledge to the production of useful materials, devices, and systems or methods, including design, development, and improvement of prototypes and new processes to meet specific requirements.”⁶³

STATEMENT OF WORK/SCOPE OF WORK: “A formal document that captures and defines the work activities, deliverables, and timeline a vendor must execute in performance of specified work for a client. The SOW usually includes detailed requirements and pricing, with standard regulatory and governance terms and conditions. “A Statement of Work is typically used when the task is well-known and can be described in specific terms. SOW’s provide explicit statements of work direction for the contractor to follow.”⁶⁴

⁵⁷ Government Accountability Office (GAO), Performance Measurement and Evaluation: Definitions and Relationships, May 2011.

⁵⁸ *Ibid.*, P. 65.

⁵⁹ *Ibid.*, P. 66.

⁶⁰ *Ibid.*, P. 66.

⁶¹ *Ibid.*, P. 67.

⁶² NIH. <https://www.nichd.nih.gov/health/clinicalresearch/clinical-researchers/steps/Pages/prepareprotocol.aspx>

⁶³ NIH Glossary of Terms. <http://grants.nih.gov/grants/glossary.htm#>

⁶⁴ General Services Administration. http://gsa.gov/graphics/fas/SOW_Application.Services.and.Component.Framework.pdf

Appendix B: Required Components of an Evaluation Report

This tool will be used by evaluators to develop final evaluation reports and assemble related documentation that will constitute appendices to the report. An evaluation report is the primary vehicle to document the methods, findings, conclusions, and recommendations of evaluations in order to disseminate results. Reports should clearly, succinctly, and impartially describe findings, conclusions, and recommendations.

Each component listed below is required for every PEPFAR evaluation report, but the order can vary and the components may be addressed in separate project documents. Agencies may have their own respective report templates/formats that should be followed, but all of the below components must be included in the final report or related documentation, which may be included as appendices to the final report. Evaluation projects that result in multiple reports on different aspects of the evaluation may be submitted as a single report, by aggregating the reports and supporting documentation. Final PEPFAR evaluation reports and related project documentation that includes information on the required components will be posted on agency websites, on pepfar.gov, as well as on a website of the Department of State, in accordance with agency open access policies.

Components	Content	Evaluation Standard Addressed (if applicable)
1. Cover and Title pages	• Title of evaluation	Standard 2
	• Date of release of report	
	• Name of evaluators and affiliation	Standard 7
2. Executive Summary	• Contains evaluation purpose, evaluation questions, brief description of project being evaluated, data collection methods, analytic methods, evaluation findings, limitations, conclusions and recommendations	
3. Project Background	• Brief description of program/project to be evaluated including dates of project implementation, total cost, geographical location, and objectives	Standard 2
4. Evaluation purpose and Questions	• Purpose of the evaluation and justification	Standard 2
	• Questions the evaluation will answer; Hypothesis (where appropriate)	
5. Evaluation Design, Methods, and Limitations	• Overall evaluation design	Standard 5
	• Type of evaluation (process, outcome, impact, economic)	Standard 3
	• Summary of stakeholder engagement	Standard 1
	• Sampling strategy	Standard 3
	• Data collection methods and rationale as aligned to evaluation questions • Sources of data Analytic methods and rationale	Standard 3

Components	Content	Evaluation Standard Addressed (if applicable)
5. Evaluation Design, Methods, and Limitations continued...	<ul style="list-style-type: none"> Ethical considerations and assurances (e.g., non-research determination and/or IRB approval with dates; application of informed consent, if appropriate; procedures to ensure human rights protection) 	Standard 3
	<ul style="list-style-type: none"> Deviations and adjustments (if any) from the approved SOW/protocol 	Standard 4
	<ul style="list-style-type: none"> Procedures used to ensure that the data are of the highest achievable quality 	Standard 8
	<ul style="list-style-type: none"> Data analysis plan 	Standard 6
	<ul style="list-style-type: none"> Limitations of the design and analytic methods 	Standard 6
6. Findings and Conclusions	<ul style="list-style-type: none"> Key findings for program improvement in relation to evaluation questions 	Standard 9
	<ul style="list-style-type: none"> Unexpected findings 	
	<ul style="list-style-type: none"> Conclusions 	
	<ul style="list-style-type: none"> Graphical representation of results and quotes where relevant 	
7. Recommendations	<ul style="list-style-type: none"> Actionable, feasible, and specific recommendations aligned to key findings 	Standards 9 and 10
8. Dissemination	<ul style="list-style-type: none"> Dissemination procedures/plan 	Standard 11
9. References	<ul style="list-style-type: none"> Reports or publications cited in the body of the report 	Standard 3
10. Appendices	<ul style="list-style-type: none"> Approved Evaluation SOW/Protocol 	Standard 6
	<ul style="list-style-type: none"> Data collection instruments/tools 	Standard 6
	<ul style="list-style-type: none"> Informed Consent, where relevant 	Standard 4
	<ul style="list-style-type: none"> Abridged bios of the evaluation team members including qualifications, experience, role on the team, and Ethical certifications (if applicable) 	Standards 7 and 4
	<ul style="list-style-type: none"> Conflict of interest statement 	Standard 7
	<ul style="list-style-type: none"> Evaluation costs 	Standard 5
	<ul style="list-style-type: none"> Project Results Framework or Logical Framework 	

Appendix C: Evaluation Resources

The below are evaluation resources and references for further information.

- The U.S. President's Emergency Plan for AIDS Relief, PEPFAR Evaluation Standards of Practice, <http://www.pepfar.gov/documents/organization/221324.pdf>
- PEPFAR Stewardship and Oversight Act of 2013 (S. 1545), <http://www.gpo.gov/fdsys/pkg/PLAW-113publ56/pdf/PLAW-113publ56.pdf>
- Department of State Program Evaluation Policy, <http://www.state.gov/s/d/rm/rls/evaluation/2012/184556.htm>
- USAID Evaluation Policy, Learning from Experience, <http://transition.usaid.gov/evaluation/USAIDEvaluationPolicy.pdf>
- USAID, Automated Directives System 203, Assessing and Learning, <http://www.usaid.gov/sites/default/files/documents/1870/203.pdf>
- CDC Evaluation Framework, <http://www.cdc.gov/eval/framework/index.htm>
- UNAIDS, 2010, Basic Terminology and Frameworks for Monitoring and Evaluation, http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/7_1-Basic-Terminology-and-Frameworks-MEF.pdf
- American Evaluation Association (AEA), Guiding Principles for Evaluators, <http://www.eval.org/p/cm/ld/fid=51>
- American Evaluation Association (AEA), An Evaluation Roadmap for a More Effective Government, September, <http://www.eval.org/EPTF/aea10.roadmap.101910.pdf>
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users (3rd ed.). Thousand Oaks, CA: Sage, <http://www.icsee.org/program-evaluation-standards-statements>
- African Evaluation Association, African Evaluation Guidelines - Standards and Norms, <http://www.afrea.org>
- Government Accountability Office (GAO), President's Emergency Plan for AIDS Relief: Agencies Can Enhance Evaluation Quality, Planning, and Dissemination, <http://www.gao.gov/products/GAO-12-673>
- Government Accountability Office (GAO), Performance Measurement and Evaluation: Definitions and Relationships, <http://www.gao.gov/new.items/d11646sp.pdf>

- Government Accountability Office, 2012, *Designing Evaluations, Revision*, GAO-12-208G. Washington, D.C. January 2012, http://www.ignet.gov/goa_588146.pdf
- Institute of Medicine (IOM), 2013. *Evaluation of PEPFAR*, The National Academies Press: Washington, DC, http://www.nap.edu/catalog.php?record_id=18256
- World Bank, *Monitoring and Evaluation: Some tools, methods, and approaches*, [http://lnweb90.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/a5efbb5d776b67d285256b1e0079c9a3/\\$FILE/MandE_tools_methods_approaches.pdf](http://lnweb90.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/a5efbb5d776b67d285256b1e0079c9a3/$FILE/MandE_tools_methods_approaches.pdf)
- The Organization for Economic Co-operation and Development (OECD). *Evaluating Development Co-operation, Summary of Key Norms and Standards*, 2nd Edition, <http://www.oecd.org/development/evaluation/dcdndep/41612905.pdf>
- Chisholm, D. and D.B. Evans (2007). Economic evaluation in health: saving money or improving care? *Journal of Medical Economics*, **10**(3): p. 325-337.
- Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, **1**(1), 1.
- World Health Organization (2008). *Framework for operations and implementation research in health and disease control programs*. http://www.who.int/hiv/pub/operational/or_framework.pdf

Appendix D: Reference Materials

(1) PROCESS EVALUATION

Definition: A type of evaluation that focuses on program/ intervention implementation, including, but not limited to access to services, whether services reach the intended population, how services are delivered, client satisfaction and perceptions about needs and services, management practices. In addition, a process evaluation might provide an understanding of cultural, socio-political, legal, and economic contexts that affect implementation of the program/intervention⁶⁵.

Example Evaluation Questions:

- To what extent were program activities/intervention(s) implemented as intended?
- Were programs implemented according to quality standards?
- To what extent was the right target population reached?
- How did socio/political/cultural factors hinder and/or facilitate program implementation?
- To what extent did capacity-building activities result in improved financial management, organizational planning, and procurement capacity, among institutions?

Timeframe: Assessment of the implementation process are conducted during implementation and/or immediately post-intervention.

Data types, sources: Interviews, focus groups, audits, observations, surveys (i.e., qualitative and quantitative).

Example Indicators: Results as compared to goals/targets, counts of clients reached, products distributed, providers trained; acceptability of services, and client satisfaction.

(2) OUTCOME EVALUATION

Definition: A type of evaluation that determines if and by how much, intervention activities or services achieved their intended outcomes⁶⁶. This form of evaluation assesses the extent to which a program achieves its outcome-oriented objectives. It focuses on outputs and outcomes (including unintended effects) to judge program effectiveness but may also assess program process to understand how outcomes are produced⁶⁷.

Example Evaluation Questions:

- To what extent did the program activities/intervention(s) reach expected short-term outcomes/changes/effects? (e.g., changes in participants' knowledge, skills, attitudes, intentions, health status)
- To what extent did the program activities/intervention reach expected intermediate-term outcomes? (e.g., changes in behaviors, norms, procedures, policies, performance at the organizational and system level)
- Did the program or intervention work, by what magnitude, and, if so, for whom?
- Were providers who received the training more likely to effectively counsel, screen and treat patients than those who did not?
- Did the program have any unintended (beneficial or adverse) effects on the target population(s)?

⁶⁵ UNAIDS, 2010, Basic Terminology, *op. cit.*, P. 65.

⁶⁶ UNAIDS, 2008, Basic Terminology and Frameworks for Monitoring and Evaluation. UNAIDS Monitoring and Evaluation Fundamentals

⁶⁷ GAO, 2011, Performance Measurement And Evaluation Definitions and Relationships

	Timeframe: Short-term, intermediate, and/or long-term post program implementation.
	Data Types, Sources: Interviews, observations, focus groups, surveys, questionnaires, clinical assessments, lab values, dose effects, etc.
	Indicators: Knowledge, risk behavior, infection status, or other endpoints, where groups are intended to serve as representative samples of the population or sub-population.

(3) ECONOMIC EVALUATION	
Definition: Use of applied analytical techniques to identify, measure, value and compare the costs and outcomes of alternative interventions. Economic evaluation is a systematic and transparent framework for assessing efficiency focusing on the economic costs and outcomes of alternative programs or interventions. This framework is based on a comparative analysis of both the costs (resources consumed) and outcomes (health, clinical, economic) of programs or interventions. Economic evaluations are also used to ensure that resources are used efficiently ⁶⁸ .	Example Evaluation Questions: <ul style="list-style-type: none"> ○ What are the costs and benefits of differentiated models of care? ○ What is the cost effectiveness of a new treatment protocol? ○ How can programs minimize service delivery costs? ○ What is the cost-effectiveness of this intervention in improving patient outcomes as compared to other treatment models?
	Timeframe: Economic evaluations are often conducted retrospectively, once a program has proven to be effective but before scale up. Sometimes, they can be done prospectively to ensure efficient allocation of resources.
	Data types, Sources: survey data, quantitative and qualitative program data on service delivery, cost and expenditure data
	Indicators: Program costs, participant costs, Quality adjusted life years (QALY), disability adjusted life year (DALY), Cost per infection averted, etc.

⁶⁸Chisholm, D. and D.B. Evans (2007). Economic evaluation in health: saving money or improving care? *Journal of Medical Economics*, **10**(3): p. 325-337.

(4) IMPACT EVALUATION

<p>Definition: Measure the change in an outcome that is attributable to a defined intervention by comparing actual impact to what would have happened in the absence of the intervention (the counterfactual scenario). Impact evaluations (IEs) are based on models of cause and effect and require a rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. There are a range of accepted approaches to applying a counterfactual analysis, though IEs in which comparisons are made between beneficiaries that are randomly assigned to either an intervention or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured to demonstrate impact⁶⁹.</p>	<p>Example Evaluation Questions:</p> <ul style="list-style-type: none"> o Does an intervention result in behavior change as a direct result of the intervention? o Does the program improve HIV or its proxy endpoint at the population-level, when properly scaled? o What are the net effects of the program in achieving long-term outcomes? (e.g., changes in prevalence, incidence, mortality, sustainability)
	<p>Timeframe: These are generally not measured accurately in the days or months post-intervention, but often long-term post implementation.</p>
	<p>Data Types, Sources: Quantitative biological and/or behavioral survey data that may or may not include recipients who directly received the intervention, Epidemiological surveillance, qualitative data, etc.</p>
	<p>Indicators: Incidence, prevalence, mortality, transmission rates, etc.</p>

⁶⁹ PEPFAR 2014 Country Operational Guidance and PEPFAR 2012 Supplemental Guidance on Implementation Science/Impact Evaluation.

(5) IMPLEMENTATION SCIENCE

<p>Definition: Implementation research is the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and to improve the quality and effectiveness of health services, in part through the study of influences on healthcare professional and organizational behavior⁷⁰.</p>	<p>Sample Research Questions:</p> <ul style="list-style-type: none"> ○ How can we best transfer scientific advances with proven efficacy and effectiveness to routine public health practice in places beyond where the approach was studied?
	<p>Timeframe: Effects assessment depends on the study question and endpoint. These may be short-term and in instances of HIV endpoints, it may be intermediate or long-term (months or years).</p>
	<p>Data types, sources: Behavioral surveys, interviews, within/between groups comparisons, audits, simulated data; mathematical modeling on cost, impact, pace and scale to guide policy and program delivery</p>
	<p>Indicators: Changes in clinical practice or services, patient experiences, or clinical/service context, etc.</p>

(6) OPERATIONS RESEARCH

<p>Definition: OR is a scientific approach to decision-making about how to design, operate, and improve programs and systems, usually under conditions requiring the allocation of scarce or finite resources. Further, it seeks to identify solutions to problems that limit program quality, efficiency and effectiveness, or to determine which alternative service delivery strategy would yield the best outcomes⁷¹.</p>	<p>Sample Research Questions:</p> <ul style="list-style-type: none"> ○ How can we improve the efficiency or effectiveness, and minimize barriers to the success, of a given program or service? ○ How can finite resources be optimized for maximum benefit of the program?
	<p>Timeframe: Effects are assessed during implementation and scale up of the intervention to monitor challenges and changes in the intervention.</p>
	<p>Data types, sources: Qualitative and quantitative data generated from brainstorming-based SWOT (strengths, weaknesses, opportunities and threats) analysis, behavioral simulation exercises, scenario analysis, system dynamics modeling and various combinations of simulation and mathematical modeling.</p>
	<p>Indicators: Evidence of program improvement, broadly defined.</p>

⁷⁰ Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, 1(1), 1.

⁷¹ World Health Organization (2008). *Framework for operations and implementation research in health and disease control programs*. http://www.who.int/hiv/pub/operational/or_framework.pdf