



UNAIDS Monitoring and Evaluation Fundamentals

A NATIONAL EVALUATION AGENDA FOR HIV

A NATIONAL EVALUATION AGENDA FOR HIV

FORWARD

Dear Colleagues,

I would like to welcome you to the UNAIDS Monitoring and Evaluation Fundamentals series. As the response to the global HIV epidemic continues to evolve, monitoring and evaluation (M&E) has become more important than ever. Determining what programs do or do not work; implementing programs with proven cost-effectiveness; monitoring progress towards achieving targets; and ensuring accountability are objectives which are especially important now in the HIV response, as well as in other health and development areas. Thus, it is increasingly important that M&E is better understood, communicated in simplified language, and conducted in a coordinated and sustainable manner that generates information that can be easily used. Further, it is essential that M&E addresses the needs of and involves all key stakeholders right from the start and that results are made publicly available and utilized strategically in policy-making, planning, and program improvement.

This series provides a common sense introduction to a range of M&E issues. It covers the fundamentals and their practical applications and includes techniques and tools for managing M&E of the HIV epidemic and response. Although the series uses HIV as its focus, the M&E fundamentals are also relevant to other areas of public health and development. As such, these books may also be useful in strengthening national M&E systems designed to track progress in other health and development goals, such as those outlined in the United Nations Millennium Development Goals (MDGs).

I hope you find this series useful and welcome your feedback and suggestions on this and future topics for the series.

With my best regards,

Deborah Rugg, PhD

Chief, UNAIDS Monitoring and Evaluation Division

ACKNOWLEDGEMENTS

This book was written by Greet Peersman (Payson Center for International Development, Tulane University, United States of America), with input from Deborah Rugg and Eva Kiwango (UNAIDS, Switzerland) and David Hales (Payson Center for International Development, Tulane University, United States of America).

Special thanks are due to the national AIDS programmes in Botswana, Mozambique, Papua New Guinea and Rwanda. Not only did they allow us to use their specific country examples, but the contents of this book are based on their experiences with moving a national evaluation agenda forward.

Thanks are also due to the following reviewers for providing valuable comments: Ansari Ameen (Caribbean Centre for Health Research, Trinidad and Tobago), Taoufik Bakkali (UNAIDS, Papua New Guinea), Boga Fidzani (National AIDS Coordinating Agency, Botswana), Wayne Gill (UNAIDS, Botswana), Elisabetta Pegurri (UNAIDS, Rwanda) and Kelvin Sikwibele (African Comprehensive HIV Partnerships, Botswana).

CONTENTS

ABBREVIATIONS	09
GETTING STARTED	11
THE FUNDAMENTALS	13
What is evaluation?	14
What is a national evaluation agenda?	19
Why is a national evaluation agenda important?	21
Who are the key stakeholders?	24
What are important evaluation standards?	27
What is the difference between evaluation and research?	29
TOOLS AND TECHNIQUES	33
What process can be used to establish a national evaluation agenda?	34
How to compile a national inventory of evaluation studies	43
What process can be used to agree on evaluation priorities?	47
How to compile a national inventory of evaluation capacity	50
How to strengthen local evaluation capacity	53
What is a costed workplan for evaluation studies?	56
How to manage implementation of evaluations	57
SUMMARY	61
QUESTIONS TO CONSIDER	65
QUICK QUIZ	69
GLOSSARY	73

LEARNING MORE ABOUT EVALUATION	79
REFERENCES	83
APPENDICES	89
Programme Evaluation Standards	90
Example checklists for research determination and institutional review board requirements	95
Example of standardized terminology to describe evaluation reports	98
Example of a structured abstract of an evaluation study	100
Example of an agenda for an evaluation priority-setting workshop	104

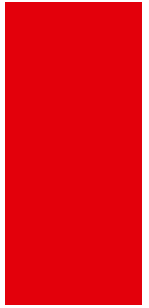
ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
AusAID	Australian Government Overseas Aid Program
CDC	Centers for Disease Control and Prevention
DFID	United Kingdom Department for International Development
Global Fund	Global Fund to Fight AIDS, Tuberculosis and Malaria
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit
HIV	human immunodeficiency virus
M&E	monitoring and evaluation
MDG	Millennium Development Goals
PEPFAR	United States President's Emergency Plan for AIDS Relief
TWG	technical working group
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNFPA	United Nations Population Fund
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WHO	World Health Organization

GETTING STARTED

The use of findings from evaluation studies is essential in guiding an effective and efficient response to the HIV epidemic. Evaluation studies can determine which programmes work and do not work, why this is the case and how to maximize available resources. Identifying evaluation priorities by means of a national, coordinated process can help to ensure that evaluations are relevant to the country's needs and that evaluators work together to avoid duplication of effort. However, a national evaluation agenda is more than just a list of prioritized evaluation questions. It is a comprehensive and standards-based approach for identifying, developing and implementing HIV evaluations and using the results to improve programmes. This book addresses the importance of evaluation for public health practice, the contents of a national evaluation agenda and the stakeholders involved, the process to establish a national evaluation agenda and the management structures to support it.

Although this book uses the area of HIV as its focus, the fundamentals, tools and techniques described are relevant to other areas of public health and social programmes.



THE FUNDAMENTALS

WHAT IS EVALUATION?

Evaluation is the systematic collection of information about the activities, characteristics and outcomes of a specific programme to determine its merit or worth. If a programme is judged to be of merit, it is also important to determine whether it is worth its cost. Evaluation provides credible information for improving programmes, identifying lessons learned and informing decisions about future resource allocation.

The investigation of any public health problem begins with asking fundamental questions:

- What is the nature of the problem, who is it affecting and what is the extent of it?
- What factors are contributing to the problem?
- What can be done to alleviate it?

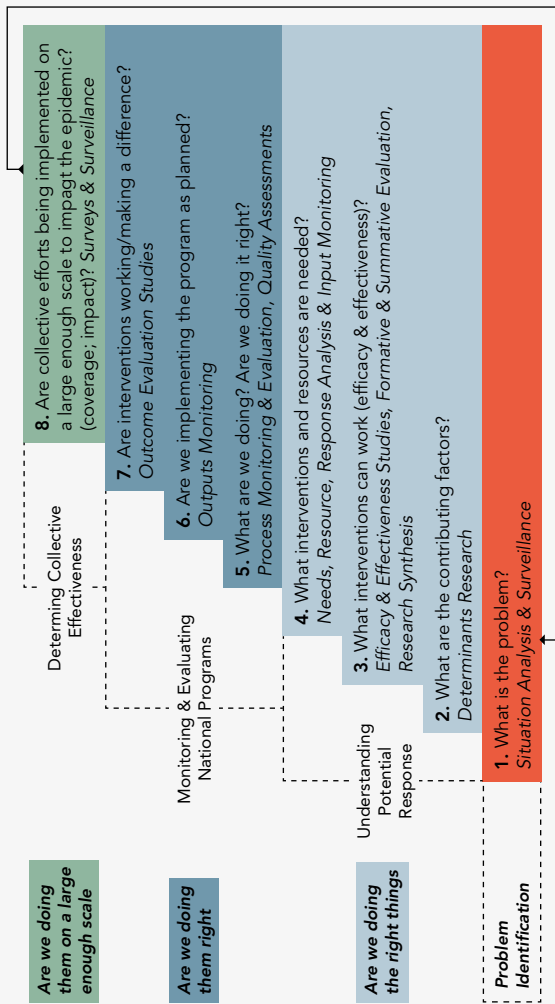
Once an appropriate programmatic response has been determined, questions are focused on:

- Is the programme working?
- Is the programme reaching enough people to reduce the impact of the problem or, ideally, eliminate it?

Figure 1 shows an overview of the key questions to ask about HIV and of the variety of data collection methods that need to be put in place to gather the right information to answer these questions. Data collection clearly needs to go beyond routine monitoring of the HIV epidemic and response. Research and evaluation play a key role in determining the three fundamental questions:

- Are we doing the right things?
- Are we doing them right?
- Are we doing them on a large enough scale to make a difference?

FIGURE 1. A PUBLIC HEALTH QUESTIONS APPROACH TO HIV MONITORING AND EVALUATION



Source: Rugg D, Carael M, Boerma T, Novak J. Global advances in monitoring and evaluation of HIV/AIDS; from AIDS case reporting to program improvement. In: Global Advances in Monitoring and Evaluation of HIV/AIDS, Rugg D, Peersman G, Carael M (Eds). New Directions for Evaluation, 2004; 103:33-48.

Programme managers can use evaluations to determine:

- Whether the programme is implemented according to agreed quality standards.
- Whether the programme is reaching its intended clients.
- Whether the clients are satisfied with the service provided.
- Whether the intended changes are occurring.

Evaluation provides the justification for which programmes need to be implemented, and how. It can answer questions of efficacy (i.e. achieving intended results under controlled conditions), effectiveness (i.e. achieving intended results under real-world conditions), efficiency (i.e. optimizing the use of limited resources) and client satisfaction (i.e. acceptability of the programme). Evaluation can give us alternatives to consider: better services or better ways of implementing services. It provides the information needed for identifying best practices and determining lessons learned.

EVALUATION BEGINS WITH ASKING THE RIGHT QUESTIONS¹

Strategy	Are we doing the right things? Providing a rationale/justification Providing a clear theory of change
Operation	Are we doing things right? Achieving intended results Optimizing limited resources Achieving client satisfaction
Learning	Are there better ways of doing things? Assessing alternatives Determining best practices Identifying lessons learned

¹ Adapted from Kusek J, Rist R. 10 Steps to a Results-based Monitoring and Evaluation System. A Handbook for Development Practitioners. Washington, DC (USA): World Bank, 2004.

WHAT IS A NATIONAL EVALUATION AGENDA?

A national evaluation agenda is a comprehensive and standards-based approach to identifying, developing and implementing HIV evaluations and using the findings to improve programmes. A national evaluation agenda contains the following elements:

1. **A national HIV evaluation strategy.** This strategy describes the rationale, goals and specific objectives for HIV evaluation and the procedures for the coordination, implementation and management of evaluations. Mechanisms for resource mobilization and for earmarking funding to different evaluation projects are an important aspect of the strategy.
2. **A process and supportive infrastructure to identify and prioritize evaluation gaps.** The process describes how evaluation needs will be identified, how often and by whom. The supportive infrastructure describes what needs to be in place² to facilitate the process and how the infrastructure is to be maintained over time.
3. **A prioritized list of evaluation questions linked to the national AIDS strategic plan.** The list includes the rationale for selecting these specific priorities, how they were selected and who was involved in the selection.
4. **A dissemination and data use strategy.** This strategy describes the key audiences for the evaluation findings, how the evaluation reports will be tailored to the different audiences and the channels through which they will be disseminated, and the mechanisms to

2 Infrastructure may include: a centralized inventory of all evaluation studies that have been conducted, planned or are ongoing in the country; a centralized inventory of existing evaluation capacity. See the Tools and Techniques section for more details.

support the use of evaluation findings for programme improvement and strategic planning.

5. **A costed operational plan.** The plan specifies the key tasks to ensure implementation of the prioritized evaluations, the actors responsible, the estimated budget requirements, the funding already secured, the timeline and the products.

The national evaluation agenda should be closely linked to the national, multisectoral, multi-year monitoring and evaluation plan and the annual monitoring and evaluation workplan. Ideally, the national evaluation agenda should be fully integrated in the national planning and implementation processes. Detailed documentation of the evaluation agenda may be an addendum to the national monitoring and evaluation plan.

WHY IS A NATIONAL EVALUATION AGENDA IMPORTANT?

The importance of evaluation to improve the effectiveness and efficiency of the national HIV response cannot be overstated. Prevention efforts are particularly important to evaluate because they continue to be woefully insufficient; for example, in 2007, for every two patients placed on antiretroviral therapy, five new HIV infections occurred. Evaluation can provide the data to support the adoption of a combination of prevention approaches, based on scientifically-derived evidence and bottom-up wisdom and ownership of local communities. Synergies between prevention, care and treatment programmes need to be built, but knowledge about which combinations of programmes work best and under what circumstances is limited. Consequently, it is essential to evaluate programmes and understand how contextual factors interact with and affect those programmes.

Not only are there still important knowledge gaps, but the AIDS community has also been criticized for being slow to implement proven approaches on an adequate scale. With existing knowledge, far more should have been, and can still be, done to curb HIV transmission globally. One problem is the lack of skills to interpret the data to guide management and to develop management structures that can effectively translate managerial decisions into actions at the point of delivery. While tackling some of these issues goes beyond the implementation of a national evaluation agenda, evaluation is a key factor in moving towards an evidence-based HIV response that is locally driven and empowered.

“Spending on assessment should not be seen as a distraction from the urgency to save lives now, but rather as an investment that has the power to save lives over the medium to long term. If the global community fails to invest now in the development of this knowledge base, in five years we will be no better informed than we are at present about which control approaches work and which do not, especially in countries with high HIV prevalence and weak health systems, where AIDS control presents the greatest challenge. Such ignorance could irreparably reverse the current global commitment towards mobilizing the necessary resources to bring about a world free of AIDS.”

Source: Bennet S, Boerma T, Brugha R. Scaling up HIV/AIDS evaluation. *Lancet*, 2006; 367:82.

A national evaluation agenda is important because it can:

- Promote evaluations with actionable results addressing the needs of the national AIDS programme.
- Improve coordination, thereby reducing duplication of effort and avoiding wasting valuable time and resources.
- Build on existing evaluation studies and promote synergies between evaluation studies.
- Draw on and strengthen local evaluation capacity.
- Leverage global, regional and local funding.
- Increase the visibility of local researchers and research institutions.
- Facilitate timely sharing of evaluation findings.
- Facilitate interpretation and application of evaluation findings to formulate policy and improve programmes.
- Provide a basis for assessing and documenting how evaluation findings have influenced local decision-making.

RATIONALE FOR SUPPORTING A NATIONAL AIDS EVALUATION AGENDA PROCESS IN MOZAMBIQUE

The national AIDS commission noted that evidence-based knowledge will lead to more effective HIV activities, which in turn will contribute to attaining the overall goal of reducing HIV prevalence in Mozambique. The following conditions underlined the importance of moving forward with a national evaluation agenda-setting process:

1. There is an increased demand for quality interventions;
2. There is an increased need for systematic documentation and dissemination of 'best practices'.
3. It is acknowledged that the evaluations conducted by academic institutions and civil society organizations are not well coordinated and do not respond clearly to the objectives in the 2nd National AIDS Strategy 2005–2009 in Mozambique. There is therefore a need to bring all evaluation activities under one umbrella in order to maximize gains and reduce repetition and inefficiencies.
4. Earmarked funding for evaluation is available from donors.

Source: Adapted from Mozambique National AIDS Commission. Concept Paper for Coordination and Implementation of the Initial Phase of the PEN II Research Sub-strategy, 2006.

WHO ARE THE KEY STAKEHOLDERS?

Given the variety of government and nongovernment actors, service delivery organizations, technical institutions, and international and donor agencies working on the HIV response in a country, it is important to ensure broad consensus and the participation of stakeholders in the evaluation agenda process. This does not mean that every organization needs to be involved in every decision or in every step of the process, but it does mean that all stakeholders are at the very least made aware and kept abreast of the process. In addition, they should all understand and have agreed on their specific role and responsibilities in its implementation.

Consider including representatives from a variety of key stakeholder groups, such as the following. The government sector, including:

- The national AIDS commission.
- The ministry of health.
- The ministry of education.
- The ministry of science and technology.
- The regional/provincial/district HIV coordinating committee.
- The local government authority responsible for health issues.

These groups are ultimately responsible for the implementation of an effective and efficient HIV response at the national and subnational levels, which should be guided by evaluation results.

HIV service delivery organizations, including:

- Civil society organizations (community-based, faith-based, networks of people living with HIV, umbrella organizations, etc.).
- Private sector organizations.

Managers and service staff are the key users of evaluation results to improve programmes and to identify key operational issues that need to be addressed. Their involvement in the national evaluation agenda process facilitates focus on relevant evaluations with actionable results.

Technical organizations, including:

- Local universities.
- Local research companies or institutions.
- Local evaluation associations.

These organizations can provide a broad perspective on appropriate evaluation methodologies, in addition to what can be learned from the evaluation literature. Their participation in the national evaluation agenda process may have important benefits, such as: enhanced visibility of local research institutions; increased access to funding for evaluation studies; and access to a professional network and infrastructure to share evaluation findings.

Bilateral and multilateral organizations and donor agencies, including:

- The Joint United Nations Programme on HIV/AIDS (UNAIDS), the United Nations Population Fund (UNFPA), the United Nations Children's Fund (UNICEF), the World Health Organization (WHO), the World Bank, etc.
- The Global Fund to Fight AIDS, Tuberculosis and Malaria (Global Fund).
- The Australian Government Overseas Aid Program (AusAID), the UK Department for International Development (DFID), Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), United States agencies (the Centers for Disease Control and Prevention (CDC), the United States President's Emergency Plan for AIDS Relief

(PEPFAR), the United States Agency for International Development (USAID)), etc.

Inclusion of these international partners can not only provide access to external technical support, where needed, but can leverage significant financial resources and global efforts to catalyse HIV evaluation.

WHAT ARE IMPORTANT EVALUATION STANDARDS?

A pioneer project to develop professional standards for programme evaluation was initiated in the United States in 1975. Its goal was to improve the evaluation of educational and training programmes in a variety of settings. Since then, these Programme Evaluation Standards have been revised and adapted by a range of national evaluation associations to make them relevant to other areas of investigation and to specific local conditions. The standards are generally acknowledged to be good practice and are routinely used in planning an evaluation, negotiating a contract to conduct an evaluation and reviewing progress during implementation of an evaluation.

The Programme Evaluation Standards address four main categories:

1. Utility Standards are intended to ensure that an evaluation will serve the information needs of the intended users.
2. Feasibility Standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic and frugal.
3. Propriety Standards are intended to ensure that an evaluation will be conducted legally, ethically and with due regard for the welfare of those involved in the evaluation, in addition to those affected by its results.
4. Accuracy Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine the worth or merit of the programme being evaluated.

See Appendix 1 for an overview of all standards.

In short, credible evaluations:

- Are impartial. They are objective and free of bias.
- Are systematic and technically adequate. They use sound methods of inquiry and follow a logical procedural model.
- Are valuable and useful. They add value to management actions such as strategy design, selection of interventions and resource allocation by providing credible strategic information in a cost-effective manner.
- Are user-owned. They include the main users of evaluation findings in the design, planning and implementation of the evaluation.
- Include feedback and dissemination. They enable findings to be used for policy formulation and programme improvements.

Source: American Evaluation Association. Guiding Principles for Evaluators. Fairhaven (USA): American Evaluation Association, 1994.

WHAT IS THE DIFFERENCE BETWEEN EVALUATION AND RESEARCH?

Before any formal study is given the go-ahead, many institutions require that a determination be made on the classification of the study as 'research' or 'non-research'. For the purpose of setting national evaluation priorities, it is not important whether a study will ultimately be classified as research or not. It is, however, an important issue in the management of the study. An evaluation study classified as 'research' requires two additional determinations: (1) does the research study involve human subjects?³ And, if so: (2) does the research study meet the criteria for exemption from review by an institutional review board for the protection of human subjects? An institutional review board will determine whether the study is indeed compliant with the regulations and will only approve the study if it is. Non-research studies are exempt from review by an institutional review board. The key distinction between research and non-research is briefly described below, while example checklists on the determination of research and institutional review board requirements are given in Appendix 2.

The distinction between research and non-research is often subtle. For many studies related to public health, research and non-research cannot be easily distinguished by looking at study characteristics, the methodological design of the study or how the participants are selected. The major difference between research and non-research lies in the 'primary intent' of the study.

The primary intent of research is to generate generalizable knowledge. This means that the intended benefits of a research study

3 A human subject is a living individual about whom an investigator conducting research obtains (a) data through intervention or interaction with the individual or (b) identifiable private information.

always extend beyond the individuals included as participants in the study; usually, the benefits can be applied to society at large. Evaluation in public health is essentially a practice to improve public health actions. The primary intent of this type of study (referred to as 'programme evaluation') is to assess the merit of and improve that programme; the evaluation is used as a management tool. The knowledge generated does not apply beyond the specific population and programme studied.

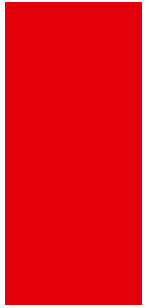
This does not imply that all evaluation studies are necessarily non-research. For example, the systematic comparison of the effectiveness of a standard programme versus an alternative programme in a study with experimental design is typically classified as research. In this case, the primary intent is to generate new knowledge and to apply this new knowledge to other settings or populations.

Note: The term 'evaluation' is used in this book to refer broadly to studies about the effectiveness, efficiency and acceptability of programmes, regardless of their classification as 'research' or 'non-research'.

KEY CHARACTERISTICS OF PROGRAMME EVALUATION AND RESEARCH

Programme evaluation			Research
Purpose	Assessing merit and worth	Improving programmes and organizations	Providing oversight/ assessing compliance
Focus	Judgement about value	Enhancement of programme services	Compliance with formal expectations
Typical mode of inquiry	Causal analysis and values inquiry	Description, with timely observation and feedback	Description, including programme activities and outcomes
<ul style="list-style-type: none"> Usual audience 	<ul style="list-style-type: none"> Organizations The public 	<ul style="list-style-type: none"> Programme managers Programme staff 	<ul style="list-style-type: none"> Legislators Funders The public
			<ul style="list-style-type: none"> Scientists 'Conventional wisdom'

Source: adapted from Mark M, Henry G, Julnes G. Evaluation. An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs. San Francisco (USA): Jossey-Bass, 2000.



TOOLS AND TECHNIQUES

WHAT PROCESS CAN BE USED TO ESTABLISH A NATIONAL EVALUATION AGENDA?

It is important to obtain buy-in from all relevant stakeholders on the need for an evaluation agenda and on the proposed process for implementation. An initial situation analysis of the local evaluation capacity (i.e. local research institutions and the studies they conducted) and of any existing procedures and oversight bodies for evaluation studies (e.g. funding mechanisms for evaluation studies, knowledge and intellectual property rights, institutional review boards, national ethics committee) provides important information to help to understand the current status and get an idea of what else needs to be done to move a national evaluation agenda forward. Agreement on the way forward needs to take advantage of already established policies and existing roles and responsibilities and draw on the comparative strengths of different stakeholders.

OVERVIEW OF KEY STEPS INVOLVED IN ESTABLISHING A NATIONAL EVALUATION AGENDA

Phase	What key tasks are involved?
Setting the stage	1. Identify key stakeholders and obtain buy-in
	2. Establish a coordination body and agree on its role
	3. Identify existing evaluation capacity
Deciding priorities	4. Compile a national inventory of relevant data/reports and completed/ongoing evaluation studies
	5. Agree priority evaluation questions
Implementing the plan	6. Develop a costed workplan
	7. Manage evaluations and disseminate the findings
	8. Use evaluation findings in decision-making

Source: Kiwango E, Sikwibele K. Process and Tools for Setting a National Evaluation/Research Agenda. Lecture, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

Since there are several steps involved in establishing a national evaluation agenda, it is useful to establish a committee or task force to coordinate the various activities. If an existing body, such as a national monitoring and evaluation technical working group, can play this role effectively there is no need to establish a new body.

Important considerations for this technical working group (committee or task force) are that it be:

- Country-driven (i.e. led by a representative of the national government).
- A neutral coordinator (i.e. has neutrality regarding the evaluation priorities).
- Multidisciplinary (i.e. includes individuals with HIV policy and programme experience as well as monitoring and evaluation experts).
- Inclusive of key stakeholders (i.e. representatives of national and subnational government sectors, major service delivery organizations from the private and civil society sectors, international agencies/donors), yet not too large in size to allow for efficient working.

AN EXAMPLE OF THE ROLES AND RESPONSIBILITIES OF A NATIONAL MONITORING AND EVALUATION WORKING GROUP: A TASK FORCE TO GUIDE THE NATIONAL EVALUATION AGENDA PROCESS IN BOTSWANA.

The national monitoring and evaluation technical working group (TWG) created an evaluation subcommittee to coordinate the national evaluation agenda process. Drawing on this existing body provided several benefits:

- The TWG was already perceived as a credible group.
- The TWG had ready access to stakeholders and relevant monitoring and evaluation data and reports.
- The TWG provides oversight for all monitoring and evaluation activities in Botswana and therefore had valuable institutional memory and a broad understanding of how different monitoring and the evaluation activities link together.

The role of the evaluation subcommittee was: (a) to lead the planning and implementation of the national evaluation agenda process; (b) to ensure country ownership and broad stakeholder buy-in and participation; and (c) to strengthen the link between evaluators and decision-makers. Specific responsibilities included:

- Advocating for the process that focused on the use of evaluations to provide data for decision-making.
- Identifying key activities, responsible organizations and timelines for implementation of the agenda-setting process.
- Mobilizing resources.
- Providing technical oversight.

- Communicating implementation progress to all stakeholders.
- Providing guidance on dissemination products for different audiences and identifying opportunities to maximize the utilization of evaluation findings.
- Supporting the development and implementation of a strategy for capacity-building in evaluation.

Source: adapted from National Evaluation Agenda Setting Guidelines. Atlanta (USA): Macro International, 2008.

It is crucial that priority setting for evaluation be evidence-based (i.e. based on data and experience). It is also important to draw on existing data to identify what is known about the effectiveness and efficiency of the HIV response. This requires an understanding of the current status of the national HIV response. A data triangulation analysis⁴ of existing data sources and summarizing the findings from evaluation studies (international and local) and how they have been applied to the country's HIV response provide important evidence for determining future evaluation needs. In addition, it is useful to know what evaluation studies are already ongoing or planned in order to avoid any duplication of effort (see below). Policy-makers and programme managers have to be consulted to identify the key policy and operational questions related to the implementation of the national strategic plan for AIDS. These need to be matched with the existing data and gaps need to be identified. A list of data gaps, which can potentially be addressed by evaluation studies

4 Data triangulation analysis can be defined as the analysis of data from three or more sources obtained by different methods. Findings can be corroborated and the weakness or bias of any of the methods or data sources can be compensated for by the strengths of another, thereby increasing the validity and reliability of the results.

(i.e. 'evaluation questions'), will provide the input for the evaluation priority setting (see below).

To ensure sustainability of a country-driven national evaluation agenda, it needs to be explicit about how opportunities for local evaluations and local evaluators will be addressed (see the Rwanda example below). Local evaluations need to draw on local expertise and contribute to local capacity-building. Historically, less than 10% of health research has been conducted in developing countries, and much of that has been undertaken by outside researchers. This has often resulted in studies that did not address the country's priority needs and/or that produced findings that were not shared locally or in a way that maximized local use for programme improvement. It is useful to compile an inventory of existing in-country capacity for evaluation (see below). This will not only inform decisions about who may be able to conduct which studies, but will also guide a capacity-building strategy for evaluation.

GUIDANCE FOR SUSTAINABLE AIDS EVALUATION IN RWANDA

- To ensure sustainable AIDS evaluation, Rwanda identified the following key actions:
- Strengthen the local evaluation infrastructure.
- Promote the use of a wide range of evaluation methods.
- Ensure evaluations are sensitive to the local culture and to gender issues.
- Promote partnerships between evaluators, service providers and local communities.

- Promote national and international collaboration between evaluators.
- Share knowledge and technologies widely.
- Maximize opportunities for local evaluators to participate in international studies.
- Strengthen dissemination mechanisms for evaluation results.

Source: Overview of Principles and Procedures of AIDS Research in Rwanda. Draft. Kigali (Rwanda): Commission Nationale de Lutte contre le SIDA, 2007.

A detailed workplan is a prerequisite for the efficient management and implementation of the selected evaluation studies (see below). The plan may also include specific activities for building evaluation capacity. The plan will then move to implementation with the selection of appropriate evaluation teams and management and data dissemination in accordance with the specifications laid out in the national evaluation strategy.

Procedures should be put in place to ensure the timely dissemination of evaluation/research findings. It is important to explore the broad range of channels for the sharing of evaluation findings, taking full advantage of current information technology. Opportunities for dissemination include: evaluation databases and websites; national/international conferences; publications in the peer-reviewed literature; a topical list serve.⁵ In Rwanda, for example, researchers are obliged to disseminate the findings to all potential users: political and administrative decision-makers, different institutions/organizations, the general public and scientific forums. Feedback from these

5 A programme that allows people to send an e-mail message to one address, which resends that message to all of the other subscribers to the list. Often, a list serve is moderated by an administrator, rather than automatically posted, to ensure that the disseminated messages are appropriate and relevant to the content area the participants subscribed to (www.taswebsites.com/info/glossary.htm).

constituents is provided back to the researchers through a committee. To ensure dissemination, the researcher/research institution needs to include a budget for the dissemination of the study results. Once the research study has been completed, a final research report is submitted. Researchers are encouraged to provide dissemination products tailored to the general public, such as bulletins and memoranda, using non-scientific language. They are also encouraged to publish their study results in local and international scientific journals and present at a variety of forums, including the on-line digital library of the national AIDS commission.⁶ The commission organizes two conferences each year for the specific purpose of sharing findings from AIDS research in Rwanda.

Last but not least, providing support for using evaluation findings for policy formulation and programme improvement is of paramount importance. It is also advisable to monitor the actual use of evaluation findings. Use of evaluation findings deserves special attention. Since evaluation reports are often presented in technical language and information on the context in which the evaluation took place is necessarily limited, the available information does not always provide a clear image of the optimum course of action to be taken. 'Translation' of evaluation findings to information that can be applied to programmes in a real-world setting requires additional effort and special talent. One limitation is the lack of skills to interpret the data to guide management and to develop management structures that can effectively translate managerial decisions into actions at the point of delivery. Apart from scarcity of capacity and weak links between researchers and policy-makers, political pressures to demonstrate positive achievements in order to secure continued funding can be a stumbling block for participation in evaluations and the use of evaluation findings. Fully understanding the context regarding the

6 Available at: <http://payson.tulane.edu/gsd/main-rwaid/main-rwaid.html>

use of HIV data, including political and cultural influences, provides important insights for determining the way forward. To ensure that the findings from evaluations are used, it is crucial to engage relevant stakeholders throughout the evaluation process: from defining relevant evaluation questions to tailored dissemination and practical application. Documenting how evaluation findings influenced planning and policy dialogues and were used in programme improvement will serve as a starting point to gather important information to determine good practices and lessons learned.

HOW TO COMPILE A NATIONAL INVENTORY OF EVALUATION STUDIES

An important first step in identifying evaluation priorities is obtaining an understanding of what is known about the effectiveness, efficiency and acceptability of HIV prevention, treatment, care and support programmes and about the important knowledge gaps. To facilitate this process, it is useful to compile a national inventory, including:

- Evaluation reports of evaluation studies conducted in the country.
- Descriptions of planned and ongoing studies in the country.

Several commercially available software packages are available to help to organize an annotated bibliography, which essentially is an evaluation inventory.

It is useful to develop a concise list of standardized terms—often referred to as ‘keywords’—that can be assigned to each document to describe its content according to general focus, specific focus and study type (see Appendix 3). Using terms that relate directly to the priority areas in the national AIDS strategic plan may provide clarity on what evaluation reports to classify under each of the programmatic areas of the plan. Other useful categories are: geographic focus (provinces, district and/or urban, rural) and target population (age, sex, specific population group). It is important that each term be clearly defined and that there is explicit guidance on how to assign the standardized terms.

The keywords can be added to the electronic register and will allow efficient retrieval of specific documents of interest from the register and a quick overview of what is contained in each document.

ELECTRONIC ANNOTATED BIBLIOGRAPHY ON HIV IN MOZAMBIQUE: DOCUMENT CHARACTERISTICS

Entry: Prevention

Title: Relationship between HIV risk perception and condom use: evidence from a population-based survey in Mozambique

Authors: Ndola P, Morris L, Mazive E, Vehidnia F, Stehr M

Publication: International Family Planning Perspectives

Date: February 2006

Location: <http://lib.bioinfo.pl>

Study type: Outcome evaluation

Target population: Youth

Geographical focus: National

Keywords: Condom distribution; information, education and communication; knowledge, attitudes, practices (KAP)

Source: adapted from Bibliografia Anotada. Estudos, Pesquisas e Documentos Relativos ao HIV/SIDA em Moçambique 1986–2007. Maputo (Mozambique): Ministry of Science and Technology, 2007.

The keywords can also be used to present a descriptive analysis of the content of the bibliography by focus areas, study types, national/international institutions involved, publication date or date of the study, completed/ongoing/planned research. This analysis can be presented as part of the background documentation for the national priority-setting workshop as it provides a concise overview of what is contained in the registry overall.

DESCRIPTIVE ANALYSIS OF EVALUATION REPORTS: CHARACTERISTICS OF HIV PREVENTION INTERVENTIONS EVALUATED (N = 142)

Characteristic N (%)

Nature of intervention

Individual behaviour component 139 (98%)
Social component 23 (16%)
Policy component 14 (10%)
Structural component 12 (8%)

Intervention components

Health education 104 (73%)
Health education only 31 (22%)
Risk-reduction supplies provision 70 (49%)
HIV counselling 41 (29%)
Service provision 30 (21%)
Skills practising 23 (16%)
Other components 37 (26%)

Intervention setting

Health 36 (25%)
Educational 26 (18%)
Commercial 19 (13%)
Community 17 (12%)
HIV 14 (10%)
Workplace 13 (9%)
Outreach 8 (6%)
Military 4 (3%)
Other setting 33 (23%)
Unspecified setting 13 (9%)

Note: figures do not add up to 142 or 100% because multiple categories may apply to each study.

Source: Peersman G, Rugg D. Intervention research and programme evaluation: the need to move beyond monitoring. In: Global Advances in HIV/AIDS Monitoring and Evaluation. Rugg D, Peersman G, Carael M (Eds). New Directions for Evaluation, 2004, 103:141–158.

If the evaluation study is formally published (e.g. in a peer-reviewed journal or in a disseminated report), it usually contains an abstract or executive summary written by the authors. However, this summary may not always contain all the relevant information to enable a quick overview of the key characteristics and findings of the evaluation study. Although time-consuming, it may be worthwhile to develop a standardized format for a structured abstract (see Appendix 4) and have a knowledgeable person extract the relevant information from the report. In this way, all studies are described similarly, which facilitates quick access to the information. It also provides an opportunity to add a critical review or other important information that may facilitate the interpretation and application of the evaluation findings.

There are several sources to access the evaluation findings of studies conducted in a variety of countries:

- The Southern African Development Community (SADC) HIV literature database: <http://www.sadc.int/sadcaidsinfo/>.
- The Cochrane Collaboration: Cochrane Collaborative Review Group on HIV Infection and AIDS: <http://www.igh.org/Cochrane/>. This website includes systematic literature reviews and meta-analysis of HIV evaluation studies.
- PEPFAR: literature digests for PEPFAR staff: <http://hivinsite.ucsf.edu/InSite?page=jl-00-00>. These digests are of recent HIV intervention research relevant to resource-limited countries. They highlight recently published studies of behavioural, policy and prevention interventions that have one or more of the following aims: to reduce sexual or drug-related risk behaviours, to decrease primary or secondary transmission, to improve health service delivery and quality of life, or to improve HIV treatment and adherence. Studies included were conducted in, or have applications to, resource-limited settings. The most novel, relevant and rigorous studies are summarized. Descriptive studies, and those less rigorous, are listed at the end of the document. Most citations are hyperlinked to PubMed. Citations for references made in the text are found (hyperlinked) at the end of each summarized study.

WHAT PROCESS CAN BE USED TO AGREE ON EVALUATION PRIORITIES?

Ideally, a national workshop is organized to discuss and provide recommendations on evaluation priorities (see Appendix 5). It is important that this event include monitoring and evaluation experts and researchers as well as policy-makers, programme managers, international agencies, key populations at higher risk and people living with HIV.

Overall, the group needs to consider where programmes and evaluations need to be targeted based on where the HIV epidemic is going: the key populations at higher risk and high transmission areas as well as other epidemic dynamics and country contexts. Participants should be thoroughly briefed on what can be learned from available monitoring and evaluation data in order to be able to make informed decisions: the cumulative knowledge from evaluations already undertaken and an analysis of the monitoring data (sero-surveillance, behavioural surveys, routine programme monitoring). Ideally, the group should be able to draw on the findings from modes of transmission studies⁷ and data triangulation analysis, including what evaluation questions are raised by assessing routine monitoring data (e.g. services to prevent mother-to-child transmission of HIV are widely available in the country, but the uptake is low; what are the reasons for this and how can service utilization be increased?).

7 Studies that estimate the size of population groups with particular risks of acquiring HIV and their exposure to HIV infection (i.e. the extent of behaviours that allow for the transmission of HIV and the prevalence of HIV infection in their contacts) to identify where most new HIV infections will occur in order to help focus interventions.

KEY PRINCIPLES TO ENSURE THE UTILITY OF HIV EVALUATIONS

Evaluations:

1. Need to be strategically focused on key issues in the local HIV epidemic and response, thereby addressing the country priorities and context.
2. Need to link explicitly to the national strategic plan for AIDS.
3. Need to be integrated in the national monitoring and evaluation plan that lays out a comprehensive monitoring and evaluation strategy and key implementation mechanisms.
4. Need to be tailored and involve stakeholders at the beginning to enhance utilization of evaluation findings at the end.
5. Need to be developed in a manner that ensures realistic and ethical implementation.

Source: Rugg D. Introduction to Evaluation. Lecture, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

Identifying key policy and operational questions linked to the national strategic plan for AIDS can be carried out best in small working groups for each major topic/programme area (e.g. prevention for youth, prevention of mother-to-child transmission of HIV, care for people living with HIV and their families). A prioritization should be determined. The results from each small group are subsequently discussed by all participants to further prioritize evaluation questions across programme areas. Ideally, not more than 10 evaluation questions should be recommended overall. The criteria used for prioritizing should be explicitly noted; for example: the programme is well-established but has never been evaluated; the programme is truly innovative and has an important potential impact, so rigorous

evaluation is warranted to assess its effectiveness before scale-up is considered. The group also needs to make a clear determination of the actions or decisions expected to be made based on the findings. This promotes utilization-focused thinking⁸ from the outset: evaluators need to design the study to obtain actionable results and programme managers need to ensure that programme conditions facilitate their application.

The prioritized evaluation questions are then considered by a small expert group of evaluators and programme managers to determine: whether data already exist to answer some of the questions; if not, what data collection methods would best be used; the anticipated time to complete the study and how this affects programme management; the feasibility and cost of collecting the data; possible evaluators; and potential financial and technical resources. Identifying priority evaluations does not preclude other evaluation studies to take place in the country, but aims to ensure that, at minimum, adequate resources are mobilized to address the priority needs. The expert group presents its findings and recommendations to the relevant national authority, which will make the final decisions on the way forward.

8 See, for example: Quinn Patton M. *Utilization-focused Evaluation*, 4th edition. Thousand Oaks (USA): Sage Publications, 2008.

HOW TO COMPILE A NATIONAL INVENTORY OF EVALUATION CAPACITY

To be able to draw on and know where to strengthen local capacity in evaluation, it is useful to assess and catalogue its current status. Such a national assessment and inventory may contain the following key elements:

- The local institutions/organizations conducting the evaluations.
- The professional background and experience of their affiliated evaluators.
- The topics focused on and types of studies undertaken.
- Their existing infrastructure to support evaluation studies.
- The national and international partnerships and collaborations the institution/organization is engaged in.
- The budget available for evaluation studies.
- The institution/organization's ability to provide relevant training or other capacity-building support.

A NATIONAL RESEARCH CAPACITY ASSESSMENT IN PAPUA NEW GUINEA

In 2008, a nationwide research⁹ capacity assessment was conducted in Papua New Guinea. Seven major research and academic institutions participated in the assessment. This assessment was followed by the development of a capacity-building plan.

Key components assessed:

- Forces in the external environment: administrative, technological, political, economic, social and cultural factors outside the institution.
- Institutional motivation: history, mission and culture.
- Institutional capacity: strategic leadership, core resources, programme and process management, interinstitutional linkages.
- Research performance: key areas related to current research performance.

Methods:

- Focus groups and one-on-one interviews with senior management, research staff, students, and technical and support staff.
- Review of publications, research project listings, annual and strategic plans of the institution.
- The specific framework for the assessment was based on a model developed by Lusthaus (1995) of the International Development Research Centre for assessing research capacity in developing countries.
- Duration: one to two weeks for each institution.

⁹ See, for example: Quinn Patton M. *Utilization-focused Evaluation*, 4th edition. Thousand Oaks (USA): Sage Publications, 2008.

Key findings:

- There was commitment and enthusiasm for research and a desire to engage in more research in the future.
- Most institutions had a research director or coordinator and a dedicated research centre with individuals conducting and publishing studies. Most had engaged in HIV-related research or in research that could be linked to issues related to HIV.
- Most institutions had an established research ethics committee.
- Each institution had representation on and actively engaged with the national AIDS commission's research advisory committee.
- Limited human resources in all institutions had an impact on the availability of research leaders who would be able to mentor others, on successfully competing for research funds and on publishing in high-ranking journals.
- The institutional infrastructures also affected the ability of researchers to engage in good-quality research. Many of the institutions did not have reliable, moderate-speed Internet access and there were limited resources to access online research databases, support fieldwork and carry out data analysis.

Conclusion:

Although research performance varied among the institutions, there is a general need to strengthen research capacity. There is a need for skills-building and training, in addition to increasing opportunities for mentorship. Interinstitutional and international collaboration should be strengthened and multidisciplinary research increased.

Source: Bakkali T. Strengthening Local Research Capacity in Papua New Guinea. Presentation, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

HOW TO STRENGTHEN LOCAL EVALUATION CAPACITY

In order to strengthen local evaluation capacity, a comprehensive approach needs to be considered that focuses not only on individual but also on institutional capacity-building. Support may include:

- Improving infrastructure to support evaluation studies, such as laboratories, information technology systems.
- Maintaining a national network of AIDS evaluators for exchange of information and sharing experiences.
- Supporting units for evaluation training in health-related institutions.
- Providing train-the-trainers courses on writing evaluation proposals to increase the ability of local evaluators to compete for local and international funding.
- Training of umbrella and affiliated organizations in evaluation methodology, research ethics and other related topics.
- Supporting visits to ethics committees and/or research institutions to exchange information and experiences.
- Providing small grants for evaluation studies.

COMPONENTS OF THE NATIONAL RESEARCH¹⁰ CAPACITY STRENGTHENING PLAN IN PAPUA NEW GUINEA

HIV research fellowship and cadet programmes (duration: five years)

Encouraging institutional twinning and collaboration.

Maximizing opportunities for mentoring, such as by experienced international evaluators, while conducting an HIV-related evaluation study.

Developing training curricula at the master's level.

Providing training in social science and HIV epidemiology at the honour's and master's levels.

The fellowship programme targets senior researchers at the doctoral and post-doctoral levels: six social scientists, two epidemiologists and two clinical scientists will be supported by means of this programme.

The cadet programme targets young researchers at the honour's and master's levels: 20 graduates will be trained in social research and 10 in HIV epidemiology to obtain the necessary skills to be actively engaged in evaluation studies.

Research capacity development fund for universities

Financial support and assistance at the strategic and coordination levels:

Building a research culture.

Strengthening research processes and increasing research quality by means of skills-building and training.

Supporting research information management and research dissemination.

Encouraging research networks and interinstitutional linkages.

HIV research skills-building training and networking

10 Research within this context refers to a wide range of studies, including evaluation studies.

Skills-building in research methods and tools for researchers currently engaged in HIV-related research.

Providing HIV research and evaluation training for health-care providers and programme implementers.

Providing opportunities for service providers and researchers to plan for research activities and share research findings.

Centre for excellence in HIV research

Maintaining a library with updated HIV information.

Developing best practice guidelines and quality assurance for HIV research.

Coordinating implementation of the national research agenda and the national research capacity strengthening plan.

Source: Bakkali T. Strengthening Local Research Capacity in Papua New Guinea. Presentation, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

WHAT IS A COSTED WORKPLAN FOR EVALUATION STUDIES?

It is important to develop a costed workplan to ensure that roles and responsibilities, deliverables and timelines are clearly identified, agreed and communicated. This plan may include:

- Activities to develop the terms of reference for each of the prioritized evaluation studies and to implement each study.
- Responsible individuals/organizations.
- Key deliverables (i.e. terms of reference, evaluation progress updates, final evaluation reports, tailored dissemination products).
- Timelines (i.e. start and end dates for activities; submission dates for deliverables).
- Financial resource requirements for all activities, including full implementation of the evaluations.
- Identified funding sources.

In addition to activities directly related to the implementation of evaluation studies, the plan may also include specific activities for building local evaluation capacity. As mentioned above, the national monitoring and evaluation workplan may already detail these activities, in which case a separate plan does not need to be prepared.

Any funding shortfalls need to be addressed as soon as possible. In fact, the workshop report and the workplan can serve as advocacy tools for leveraging additional funding. Several donor agencies, including the Global Fund to Fight AIDS, Tuberculosis and Malaria and the PEPFAR initiative, provide financial resources to support in-country evaluations. Including representatives from multilateral and bilateral agencies in the evaluation agenda process will facilitate leveraging their financial and technical resources. It is important that the workplan be shared with all relevant stakeholders and used to monitor progress on implementation of the activities.

HOW TO MANAGE IMPLEMENTATION OF EVALUATIONS

Clear, well-written terms of reference are a critical first step towards ensuring that a robust evaluation is conducted. The terms of reference are not only used to facilitate management of the evaluation planning, implementation and dissemination process but can also be used to share key information about the evaluation study with relevant stakeholders and to extract the necessary information for the national inventory of evaluation studies.

A full discussion of the management of evaluations is outside the scope of this book; references to relevant publications can be found below. However, a few of the key steps to successfully manage an evaluation are the following.

Once the evaluation workplan is agreed, the following steps need to be undertaken:

- Obtain financial and other resources.
- Advertise a tender for evaluation proposals.
- Review and select evaluation proposals.
- Select an appropriate evaluator/evaluation team that will design and implement the evaluation.
- Submit the evaluation proposal to the relevant committees for determination of research/non-research and necessary ethical procedures.
- Monitor progress of the implementation of the evaluation.
- Critically review and discuss the evaluation results with the evaluator/evaluation team.
- Ensure timely dissemination of the evaluation report and disseminate products tailored to different audiences.
- Support and document the use of the evaluation findings.

TERMS OF REFERENCE FOR AN EVALUATION STUDY

Terms of reference should include:

- A description of the evaluation topic.
 - Background information/context and rationale for undertaking the evaluation.
 - Specific objectives of the evaluation.
 - Who will use the evaluation findings and how they will be used to strengthen policies/programmes.
 - Key questions the evaluation will address.
 - The target groups addressed by the evaluation.
 - The programmes to be evaluated.
 - The data to be collected.
 - The evaluation design.
 - Key data sources and data collection procedures.
 - Key data analysis procedures.
 - The sequence and time schedule of evaluation activities.
 - The evaluation team members, their roles/responsibilities and estimated level of effort.
-
- Administrative and logistical support.
 - The evaluation budget.

Source: adapted from UNAIDS. Prevention Evaluation—Standards for Developing Terms of Reference. Geneva (Switzerland): UNAIDS, 2009.

The example from Rwanda provides some key roles and responsibilities of local institutions in managing evaluations.

ROLES AND RESPONSIBILITIES OF DIFFERENT AGENCIES IN THE IMPLEMENTATION OF AIDS EVALUATIONS IN RWANDA

The role of the national research committee

This committee includes representatives from public, private and community-based organizations involved in AIDS service provision and AIDS research. It is chaired by the national AIDS commission and meets once a month. The role of the committee includes:

- Approving evaluation proposals (as the first step in the approval process). Approval requires that the proposed study:
 - Responds to the needs of the national strategic plan for AIDS.
 - Does not present an unnecessary duplication of effort.
 - Meets quality standards to ensure that its findings are valid and useable.
 - Improves local research capacity.
- Promoting collaboration and networking between local and international researchers.
- Monitoring progress in implementation of the evaluation project.
- Ensuring that the evaluation findings are submitted to the national AIDS commission for inclusion in the national evaluation inventory and supporting their use for policy formulation and programme improvement.

The role of the national ethics committee

This committee ensures that each research study meets the ethical requirements and Rwandan law. It meets once a month to review evaluation protocols and communicates its decisions to the Ministry of Health. The role of the committee includes:

Approving evaluation proposals (as the second step in the approval process). Approval requires that the proposed study:

Meets ethical requirements throughout all phases of the research project in order to ensure that the dignity and well-being of study participants are guaranteed.

Uses explicit consent procedures and assures confidentiality of personal information.

Provides services to study participants with health problems/other needs during the course of the study.

Pays attention to the development needs of the country.

Regularly evaluating the research project for adherence with the ethical requirements.

The role of the National Institute of Statistics of Rwanda (Institut des Statistiques du Rwanda, INSR)

The National Institute of Statistics provides leadership in improving capacity to analyse and use data in decision-making. The role of the Institute includes:

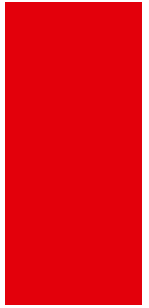
Validating the methods described in the evaluation proposals.

Developing capacity for data analysis using appropriate statistical methods.

Establishing innovative systems for data exchange and dissemination, including evaluation findings.

Strengthening evaluation capacity of high-level institutions such as universities and the Treatment and Research AIDS Centre.

Source: Overview of Principles and Procedures of AIDS Research in Rwanda. Draft. Kigali (Rwanda): Commission Nationale de Lutte contre le SIDA, 2007.

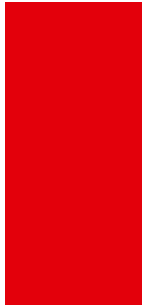


SUMMARY

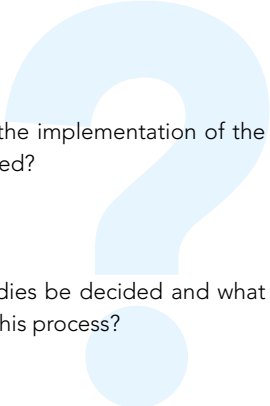
Findings from evaluation studies are essential to guide an effective and efficient response to the HIV epidemic. A national evaluation agenda is a comprehensive and standards-based approach for identifying, developing and implementing HIV evaluations and using the results for improving programmes.

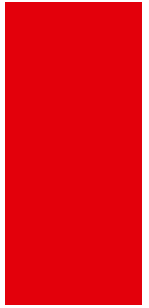
Using a coordinated process including broad consensus and the participation of relevant stakeholders can help:

- To ensure actionable results.
- To avoid duplication of effort.
- To build local capacity.



QUESTIONS TO CONSIDER

- 
- What is contained in a national evaluation agenda and why is it important to have one?
 - How can buy-in for the process and the implementation of the national evaluation agenda be obtained?
 - How can priorities for evaluation studies be decided and what evidence should be used to support this process?
 - What is the difference between evaluation and research and what is its importance for the national evaluation agenda?
 - How can evaluations draw on and support strengthening the capacity of local evaluators and institutions?
 - What system needs to be in place to support the implementation of evaluation studies?
 - What are the key issues in disseminating and using evaluation findings?



QUICK QUIZ

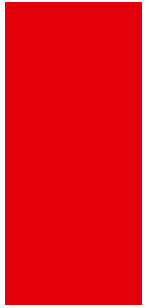
1. Indicate true or false:

- _____ Research intends to generate or contribute to generalizable knowledge to improve public health practice.
- _____ The intended benefits of a research study are applicable to the study participants only.
- _____ Evaluation is the systematic collection of information about programme activities, characteristics and outcomes that determines the merit or worth of a specific programme for use in improving programmes.
- _____ The knowledge generated through a programme evaluation is applicable beyond the population or programme from which the data were collected.

2. Why is it important to set a national agenda for HIV evaluation/research studies?

Tick all that apply:

- To understand who is doing which studies and what the existing capacity for evaluation/research is.
- To reduce proliferation and duplication of studies and waste of resources.
- To coordinate in-country resources for evaluation/research.
- To improve dissemination and sharing of evaluation/research findings.
- To improve visibility of evaluation/research studies at the national, regional and international levels.
- To monitor how evaluation/research studies have influenced programme planning and policy.



GLOSSARY

Accountability. Responsibility for the use of resources and the decisions made, as well as the obligation to demonstrate that work has been done in compliance with agreed-upon rules and standards and to report fairly and accurately on performance results vis-a-vis mandated roles and/or plans.

Assumptions. Hypotheses about factors or risks which could affect the progress or success of an intervention. Intervention results depend on whether or not the assumptions made, prove to be correct.

Attribution. The ascription of a causal link between observed changes and a specific intervention.

Beneficiaries. The individuals, groups, or organizations, whether targeted or not, that benefit directly or indirectly, from the intervention.

Economic evaluation. Use of applied analytical techniques to identify, measure, value and compare the costs and outcomes of alternative interventions. Types of economic evaluations include cost-benefit, cost-effectiveness, cost-efficiency evaluations.

Effectiveness. The extent to which a programme or intervention has achieved its objectives under normal conditions in a real-life setting.

Efficacy. The extent to which an intervention produces the expected results under ideal conditions in a controlled environment.

Efficiency. A measure of how economically inputs (resources such as funds, expertise, time) are converted into results.

Evaluability. Extent to which an intervention or programme can be evaluated in a reliable and credible fashion.

Evaluation. The rigorous, scientifically-based collection and analysis of information about programme or intervention activities, characteristics, and outcomes that determine the merit or worth of the programme or intervention. Evaluation studies provide credible information for use in improving programmes or interventions, identifying lessons learned, and informing decisions about future resource allocation.

Experimental evaluation design. An evaluation design that compares observations in participants who are randomly assigned to a programme group (i.e., the experimental group) with those that are randomly assigned to a control group.

Formative evaluation. A type of evaluation intended to improve the performance of a programme or intervention. A formative evaluation is usually undertaken during the design and pretesting of the intervention or programme, but it can also be conducted early in the implementation phase, particularly if implementation activities are not going as expected.

Generalizability. The extent to which findings can be assumed to be true for the entire target population, not just the sample of the population under study. Note: To ensure generalizability, the sampling procedure and the data collected need to meet certain methodological standards.

Impact. The long-term, cumulative effect of programmes or 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 programme or intervention, but a specific programme or intervention may, together with other programmes or interventions, contribute to impacts on a population.

Impact evaluation. A type of evaluation that assesses the rise and fall of impacts, such as disease prevalence and incidence, as a function of HIV programmes or interventions. Impacts on a population seldom can be attributed to a single programme or intervention; therefore, an evaluation of impacts on a population generally entails a rigorous design that assesses the combined effects of a number of programmes or interventions for at-risk populations.

Internal evaluation. An evaluation of an intervention conducted by a unit and/or individuals who report to the management of the organization responsible for the financial support, design and/or implementation of the intervention.

Meta-evaluation. A type of evaluation designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators.

Operational research. Systematic and objective assessment of the availability, accessibility, quality, and/or sustainability of services designed to improve service delivery. It assesses only factors that are under the control of programme or project managers, such as improving the quality of services, increasing training and supervision of staff members, and adding new service components.

Outcome. Short-term and medium-term effect of an intervention's outputs, such as change in knowledge, attitudes, beliefs, behaviours.

Outcome evaluation. A type of evaluation that determines if, and by how much, intervention activities or services achieved their intended outcomes. An outcome evaluation attempts to attribute observed changes to the intervention tested. Note: An outcome evaluation is methodologically rigorous and generally requires a comparative

element in its design, such as a control or comparison group, although it is possible to use statistical techniques in some instances when control or comparison groups are not available (e.g., for the evaluation of a national programme).

Process evaluation. A type of evaluation that focuses on programme 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 contexts that affect implementation of the programme or intervention.

Programme evaluation. A study that intends to control a health problem or improve a public health programme or service. The intended benefits of the programme are primarily or exclusively for the study participants or the study participants' community (i.e., the population from which the study participants were sampled); data collected are needed to assess and/or improve the programme or service, and/or the health of the study participants or the study participants' community. Knowledge that is generated does not typically extend beyond the population or programme from which data are collected.

Relevance. The extent to which the objectives, outputs, or outcomes of an intervention are consistent with beneficiaries' requirements, organizations' policies, country needs, and/or global priorities.

Reliability. Consistency or dependability of data collected through the repeated use of a scientific instrument or a data collection procedure used under the same conditions.

Research. A study which intends to generate or contribute to generalizable knowledge to improve public health practice, i.e., the study intends to generate new information that has relevance beyond the population or programme from which data are collected. Research typically attempts to make statements about how the different variables under study, in controlled circumstances, affect one another at a given point in time.

Stakeholder. A person, group, or entity who has a direct or indirect role and interest in the goals or objectives and implementation of a programme or intervention and/or its evaluation.

Summative evaluation. A type of evaluation conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. It is designed to provide information about the merit or worth of the intervention.

Target group. Specific group of people who are to benefit from the result of the intervention.

Terms of reference (of an evaluation). Written document presenting the purpose and scope of the evaluation, the methods to be used, the standards against which performance is to be assessed or analyses to be conducted, the resources and time allocated, and the reporting requirements.

Validity. The extent to which a measurement or test accurately measures what is intended to be measured.



LEARNING MORE ABOUT EVALUATION

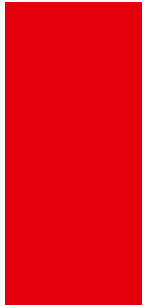
A large number of reference documents are available on evaluation methods and procedures.

Some useful resources:

- Centers for Disease Control and Prevention. Framework for program evaluation in public health. *MMWR*, 1999; 48 (No. RR-11):1–40.
- Mark M, Henry G, Julnes G. Evaluation. *An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco (USA): Jossey-Bass, 2000.
- Quinn Patton M. *Utilization-focused Evaluation*, 4th edition. Newberg Park (USA): Sage Publications, 2008.
- Rossi P, Freeman H. *Evaluation: a Systematic Approach*. Newberg Park (USA): Sage Publications, 1993.

Some useful websites:

- African Evaluation Association (available at: <http://www.afrea.org/home/index.cfm>)
- American Evaluation Association (available at: <http://www.eval.org>)
- Australasian Evaluation Association (available at: <http://www.aes.asn.au>)
- Evaluation Associations Around the World (available at: http://www.ioce.net/members/eval_associations.shtml)
- European Evaluation Society (available at: <http://www.europe-anevaluation.org>)
- Organisation for Economic Development and Cooperation/Development Cooperation Directorate (available at: https://www.oecd.org/department/0,3355,en_2649_34435_1_1_1_1_1,00.html)
- Planning, Monitoring, Evaluation, and Systematisation of Anti-rural Poverty Projects in Latin America and the Caribbean (available at: <http://www.preval.org/>)
- United Nations Interagency Working Group (available at: <http://www.uneval.org>)



REFERENCES

American Evaluation Association. *Guiding Principles for Evaluators*. Fairhaven (USA): American Evaluation Association, 1994.

American Evaluation Association. *The Program Evaluation Standards. Summary of the Standards*. <http://www.eval.org/EvaluationDocuments/progeval.htm>, accessed on 30 March 2009.

Bakkali T. *Strengthening Local Research Capacity in Papua New Guinea*. Presentation, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

Bennett S, Boerma JT, Brugha R. Scaling up HIV/AIDS evaluation. *Lancet*, 2006; 367:79–82.

Bertozzi S, Laga M, Bautista-Arredona S, Coutinho A. Making HIV prevention programmes work. *Lancet*, 2008; 372:831–844.

Centers for Disease Control and Prevention. Framework for program evaluation in public health. *MMWR*, 1999; 48 (No. RR-11):1–40.

Centers for Disease Control and Prevention & Agency for Toxic Substances and Disease Registry. *Procedures for Protection of Human Research Participants*, 2003. <http://www.cdc.gov/od/ads/hsr2.htm>, accessed on 31 January 2009.

Commission Nationale de Lutte contre le SIDA. *Overview of Principles and Procedures of AIDS Research in Rwanda. Draft*. Kigali (Rwanda): Commission Nationale de Lutte contre le SIDA, 2007.

Fink A. *Evaluation Fundamentals*. Newberg Park (USA): Sage Publication Inc, 1993.

Global Fund to Fight AIDS, Tuberculosis and Malaria. *Framework for Operations and Implementation Research in Health and Disease Control Programmes*. Geneva (Switzerland): Global Fund, 2008.

Kiwango E, Sikwibele K. *Process and Tools for Setting a National Evaluation/Research Agenda*. Lecture, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

Kusek J, Rist R. *Ten Steps to a Results-based Monitoring and Evaluation System. A Handbook for Development Practitioners*. Washington, DC (USA): World Bank, 2004.

Lusthaus C, Anderson G, Murphy E. *Institutional Assessment. A Framework for Strengthening Organizational Capacity for IDRC's Research Partners*. Ottawa (Canada): International Development Research Centre, 1995.

Macro International. *National Evaluation Agenda Setting Guidelines*. Atlanta (USA): Macro International, 2008.

Mark M, Henry G, Julnes G. *Evaluation. An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco (USA): Jossey-Bass, 2000.

Merson M. The history and challenge of HIV prevention. *Lancet*, 2008; 372:475–488.

Ministry of Science and Technology. *Bibliografia Anotada. Estudos, Pesquisas e Documentos Relativos ao HIV/SIDA em Moçambique 1986–2007*. Maputo (Mozambique): Ministry of Science and Technology, 2007.

National AIDS Commission. *Concept Paper for Coordination and Implementation of the Initial Phase of the PEN II Research Sub-strategy*. Maputo (Mozambique): National AIDS Commission, 2006.

National AIDS Council. *National Research Agenda for HIV and AIDS in Papua New Guinea 2008–2013*. Port Moresby (Papua New Guinea): National AIDS Council, 2008.

Peersman G, Rugg D. Intervention research and program evaluation: the need to move beyond monitoring. In: *Global Advances in HIV/AIDS Monitoring and Evaluation*. Rugg D, Peersman G, Carael M (Eds). *New Directions for Evaluation*, 2004; 103:141–158.

Rossi P, Freeman H. *Evaluation: a Systematic Approach*. Newberg Park (USA): Sage Publications Inc, 1993.

Rugg D, Carael M, Boerma T, Novak J. Global advances in monitoring and evaluation of HIV/AIDS: from AIDS case reporting to program improvement. In: *Global Advances in Monitoring and Evaluation of HIV/AIDS*, Rugg D, Peersman G, Carael M (Eds). *New Directions for Evaluation*, 2004; 103:33–48.

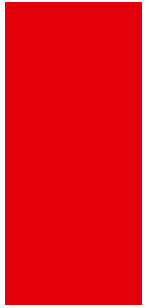
Rugg D. *Introduction to Evaluation*. Lecture, UNAIDS 3rd Global Monitoring and Evaluation Training, Bangkok (Thailand), 2008.

UNAIDS. *Organizing Framework for a Functional National HIV Monitoring and Evaluation System*. Geneva (Switzerland): UNAIDS, 2008.

UNAIDS. *Glossary of Monitoring and Evaluation Terms*. Geneva (Switzerland): UNAIDS, 2008.

UNAIDS. *Prevention Evaluation. Standards for Developing Terms of Reference*. Geneva (Switzerland): UNAIDS, 2009.

Wilson D, Halperin D (2008). "Know your epidemic, know your response": a useful approach, if we get it right. *Lancet*, 2008; 372:423-426.



APPENDICES

APPENDIX ONE: PROGRAMME EVALUATION STANDARDS

The Programme Evaluation Standards¹¹ are the following:

1. Utility (U) Standards are intended to ensure that an evaluation will serve the information needs of intended users:

- **U1. Stakeholder identification.** Persons involved in or affected by the evaluation should be identified, so that their needs can be addressed.
- **U2. Evaluator credibility.** The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance.
- **U3. Information scope and selection.** Information collected should be broadly selected to address pertinent questions about the programme and be responsive to the needs and interests of clients and other specified stakeholders.
- **U4. Values identification.** The perspectives, procedures and rationale used to interpret the findings should be carefully described, so that the bases for value judgements are clear.
- **U5. Report clarity.** Evaluation reports should clearly describe the programme being evaluated, including its context and the purposes, procedures and findings of the evaluation, so that essential information is provided and easily understood.
- **U6. Report timelines and dissemination.** Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a timely fashion.
- **U7. Evaluation impact.** Evaluations should be planned, conducted and reported in ways that encourage follow-through by

¹¹ American Evaluation Association. The Program Evaluation Standards. Summary of the Standards. Available at: <http://www.eval.org/EvaluationDocuments/progeval.html>, accessed on 30 March 2009.

stakeholders, so that the likelihood that the evaluation will be used is increased.

- **2. Feasibility (F) Standards** are intended to ensure that an evaluation will be realistic, prudent, diplomatic and frugal:
- **F1. Practical procedures.** The evaluation procedures should be practical, to keep disruption to a minimum while needed information is obtained.
- **F2. Political viability.** The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.
- **F3. Cost effectiveness.** The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified.

3. Propriety (P) Standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results:

- **P1. Service orientation.** Evaluations should be designed to assist organizations to address and effectively serve the needs of the full range of targeted participants.
- **P2. Formal agreements.** Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.

- **P3. Rights of human subjects.** Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.
- **P4. Human interactions.** Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed.
- **P5. Complete and fair assessment.** The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the programme being evaluated, so that strengths can be built upon and problem areas addressed.
- **P6. Disclosure of findings.** The formal parties to an evaluation should ensure that the full set of evaluation findings, along with pertinent limitations, are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results.
- **P7. Conflict of interest.** Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.
- **P8. Fiscal responsibility.** The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate.

4. Accuracy (A) Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine the worth or merit of the programme being evaluated:

- **A1. Programme documentation.** The programme being evaluated should be described and documented clearly and accurately, so that the programme is clearly identified.

- **A2. Context analysis.** The context in which the programme exists should be examined in enough detail, so that its likely influences on the programme can be identified.
- **A3. Described purposes and procedures.** The purposes and procedures of the evaluation should be monitored and described in enough detail, so that they can be identified and assessed.
- **A4. Defensible information sources.** The sources of information used in a programme evaluation should be described in enough detail, so that the adequacy of the information can be assessed.
- **A5. Valid information.** The information-gathering procedures should be chosen or developed and then implemented so that they will assure that the interpretation arrived at is valid for the intended use.
- **A6. Reliable information.** The information-gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable for the intended use.
- **A7. Systematic information.** The information collected, processed and reported in an evaluation should be systematically reviewed and any errors found should be corrected.
- **A8. Analysis of quantitative information.** Quantitative information in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.
- **A9. Analysis of qualitative information.** Qualitative information in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.
- **A10. Justified conclusions.** The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them.
- **A11. Impartial reporting.** Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect the evaluation findings.

- **A12. Meta-evaluation.** The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.

APPENDIX TWO: EXAMPLE CHECKLISTS FOR RESEARCH DETERMINATION AND INSTITUTIONAL REVIEW BOARD REQUIREMENTS

Before any formal study is given the go-ahead, many institutions require that a determination be made on the classification of the study as 'research' or 'non-research'. The following checklist provides concise criteria for study classifications based on the guidelines established by the United States Centers for Disease Control and Prevention. This is provided for illustrative purposes only, as regulations for research determinations are institution-specific.

CHECKLIST FOR THE DETERMINATION OF A STUDY AS 'NON-RESEARCH' OR 'RESEARCH'

The study is not research. The primary intent is improving public health practice or disease control.

- Epidemic/endemic disease control activity: collected data directly relate to disease control needs.
- Routine disease surveillance activity: data used for disease control programme or policy purposes.
- Programme evaluation activity: data are used primarily for that purpose.
- Post-marketing surveillance of efficacy and/or adverse effects of a new regimen, drug or device.
- The activity is purely administrative (e.g. purchase orders or contracts for services or equipment) and not related to research.

The study is research. The primary intent is generating generalizable knowledge.

- None of the above descriptions apply to the study.

Source: Centers for Disease Control and Prevention & Agency for Toxic Substances and Disease Registry. Procedures for Protection of Human Research Participants. Atlanta (USA), 2003. <http://www.cdc.gov/od/ads/hsr2.htm>, accessed on 31 January 2009.

The determination of an evaluation study as research has important implications: two additional determinations must be made: (1) does the research study involve human subjects?¹²; and, if so: (2) does the research study meet the criteria for exemption from review by an institutional review board for the protection of human subjects? A concise overview of such determination is provided in the following checklist. An institutional review board will determine if the study is indeed compliant with the regulations and will only approve the study if it is. Non-research studies are exempt from institutional review board review.

12A human subject is a living individual about whom an investigator conducting research obtains (a) data through intervention or interaction with the individual or (b) identifiable private information. Intervention includes both physical procedures by which data are gathered and manipulations of the subjects. Interaction includes communication or interpersonal contact between the investigator and the subject. Private information includes information about behaviour that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information that has been provided for specific purposes and that the individual can reasonably expect will not be made public. Private information must be individually identifiable (i.e. the identity of the subject is or may readily be ascertained) in order to classify the study as research involving human subjects.

CHECKLIST FOR THE DETERMINATION OF INSTITUTIONAL REVIEW BOARD REQUIREMENTS OR EXEMPTION

The study is classified as research but does not involve identifiable human subjects—the study is exempt from institutional review board review if:

- The research involves collection/analysis of data about health facilities or other units that are not individual persons; or
- The research involves data and/or specimens from deceased persons; or
- The research uses unlinked anonymous data or specimens; that is, all of the following are required: (1) no contact with human subjects is involved for the proposed activity; and (2) data or specimens are/were collected for another purpose; and (3) no extra data/specimens are/were collected for this purpose; and (4) identifying information either was not obtained or has been removed so that data cannot be linked or re-linked with identifiable human subjects.

The study is classified as research and involves identifiable human subjects—institutional review board review is required if none of the above descriptions apply to the study.

Source: Centers for Disease Control and Prevention & Agency for Toxic Substances and Disease Registry. Procedures for Protection of Human Research Participants. Atlanta (USA), 2003. <http://www.cdc.gov/od/ads/hsr2.htm>, accessed on 31 January 2009.

APPENDIX THREE: EXAMPLE OF STANDARDIZED TERMINOLOGY TO DESCRIBE EVALUATION REPORTS

A. Focus of the report / study

Specific focus: for HIV prevention-related reports / studies (i.e. studies related to preventing new HIV infections/reducing the number of new HIV infections) assign all terms that apply:

- 1.1 Activism
- 1.2 Biosafety
- 1.3 Blood safety
- 1.4 Communication strategies (includes mass media)
- 1.5 Condom distribution and promotion
- 1.6 Domestic and sexual violence
- 1.7 Formal education (includes education curricula)
- 1.8 Gender disparities
- 1.9 HIV testing and counselling
- 1.10 Information, education and communication
- 1.11 Male circumcision
- 1.12 Prevention of mother-to-child transmission
- 1.13 Sexual and reproductive health promotion
- 1.14 Sexually transmitted infections—diagnosis, treatment and control
- 1.15 Traditional practices
- 1.16 Youth-friendly services
- 1.17 Other (specify): _____ (write in)

B. Target population / study participants

1. Age: assign all terms that apply:
 - 1.1 Adults (aged 25 and older)
 - 1.2 Children (aged 1–14)
 - 1.3 Infants (birth to 1 year)
 - 1.4 Youth (aged 15–24)

2. Sex: assign all terms that apply:

Men and women

Men only

Women only

3. Population group: assign all terms that apply:

Discordant couples

General population

Health care workers

Most-at-risk: clients of sex workers

Most-at-risk: injecting drug users

Most-at-risk: men who have sex with men

Most-at-risk: mobile populations

Most-at-risk: prisoners

Most-at-risk: sex workers

Most-at-risk: uniformed services

Nongovernmental organizations

Orphans and vulnerable children

Patients—sexually transmitted infections

Patients—tuberculosis

People living with HIV

Pregnant women

Students—primary education

Students—secondary education

Students—tertiary education

Other (specify): _____ (write in)

Source: selected from standardized terminology developed for the national evaluation inventory in Mozambique. Note that the full definitions of terms are not included here.

APPENDIX FOUR: EXAMPLE OF A STRUCTURED ABSTRACT OF AN EVALUATION STUDY

ANTIRETROVIRAL THERAPY

Mills EJ, Nachega JB, Buchan I, Orbinski J, Attaran A, Singh S, et al. Adherence to antiretroviral therapy in sub-Saharan Africa and North America: a meta-analysis. JAMA 2006; 296(6): 679-90.

OBJECTIVE: This is a systematic review comparing levels of adherence to antiretrovirals among patients in sub-Saharan Africa to those who are part of more established treatment programmes in North America.

STUDIES: Fifty-eight prospective studies assessing adherence rates as a primary or secondary outcome among general population samples of HIV-positive people in North America and Africa were included: 31 from North America (28 full-text articles, three abstracts) and 27 from sub-Saharan Africa (nine full text articles, 18 abstracts). The following were excluded: studies not representative of the "general population" of HIV-positive individuals, meaning those that assessed only men who have sex with men, the homeless, children or drug users, those containing experimental adherence interventions, those presenting adherence as the mean of all doses taken. Studies measured "adherence" differently. Patient self-report was used to assess adherence in 71% of the North American and 66% of the sub-Saharan studies. Other measures included pharmacy claims, medical records, MEMS-cap (pill-counting devices), and clinician and staff reports. Time periods for measurement varied greatly, from the previous 24 hours up to 365 days, and including the previous three

days, four days, one week, two weeks, one month, two months and three months.

The definition of “being adherent” varied by study, based on an “adherence threshold,” from 100% adherence in some studies to >80% in others. For North American studies, 15/31 (48%) considered 100% adherence as “adherent”; 9/31 (29%) studies ≥95% adherence; 4/31 (13%) studies ≥90% adherence; 3/31 (10%) studies ≥80% adherence. The number of studies from Africa using each adherence threshold was the following: 11/27 (41%) studies ≥100% adherence; 11/27 (41%) studies ≥95% adherence; 2/27 (7%) studies ≥90% adherence; 3/27 (11%) studies ≥80% adherence.

SEARCH STRATEGY: Eleven electronic databases (Medline, EMBASE, Cochrane CENTRAL, AIDSLINE, AMED, CINAHL, TOXNET, Development and Reproductive Toxicology, Hazardous Substance Databank, PsycINFO, and Web of Science), major HIV conference abstract databases, lay publications and websites were searched. Individual clinical researchers, AIDS cohort trial groups, and treatment advocacy groups were contacted to inquire about unpublished studies. Searches were conducted from the inception of the particular database until 18 April 2006.

PARTICIPANTS: This review included studies of general (male and female) HIV-positive populations in North America and Africa. The median number of participants per study from North America was 219 (total=17573 patients), and from Africa was 100 (total=12116 patients).
INTERVENTION(S): None.

OUTCOME MEASURES: This review provided pooled and comparative analyses of the proportion of people who were “adherent” across studies, stratified by data source (abstract or full-text), continent

(North America or Africa), and adherence threshold used (100% to >80%). Multivariable random effects logistic regression was used to determine independent predictors of adherence; predictor variables included continent, adherence threshold used, clinic setting, and whether treatment was free or paid.

RESULTS: Meta-analysis revealed significant heterogeneity across studies. The combined analysis of data from all studies revealed adherence of 64% (95% CI, 59%-70%), meaning that 64% of persons were adherent, based on whatever definition (adherence threshold) was used in each study. Pooled estimates from North American studies revealed 55% of persons studied were adherent (95% CI, 49%-62%). A pooled adherence estimate from Africa studies showed 77% of persons were adherent (95% CI, 68%-85%). Independent predictors of adherence included African study, adherence threshold used (100% and 95%), using more than one adherence measure, and using MEMS.

CONCLUSIONS: The authors concluded that favourable levels of adherence can be achieved in sub-Saharan Africa. They state that concerns about suboptimal adherence are not supported by the data, and should not be used as a reason to delay access to treatment. However, these higher levels of adherence may be due in part to African individuals being on less complicated treatment regimens, and being early in treatment. The initial experience of dramatic clinical improvement without the long-term side effects could be related to better adherence. Thus, these findings may not persist as access becomes more widespread.

QUALITY RATING: Based on the QUOROM grading system for systematic reviews, this analysis was of high quality. The main limitation of this analysis was the quality of the studies. There was a very wide

range in how adherence was measured, little validation of measures, and many studies from Africa did not provide complete data (they were obtained from abstracts).

IN CONTEXT (Reviewer comment): This is the first meta-analysis of adherence data from sub-Saharan Africa, and comparing it to data from the United States and Canada. There has been no evidence to indicate that adherence in resource-poor settings is significantly worse than in regions where antiretroviral therapy is widely available or healthcare infrastructure more established. Data from homeless HIV-positive people in San Francisco indicate that even though adherence may be only 80%, there may be clinical improvement. This review and meta-analysis excluded the vast majority of those high-risk people (such as men who have sex with men and injecting drug users) who are on antiretroviral therapy in North America. The population of heterosexual or mixed gender populations may or may not be more similar to African populations.

PROGRAMMATIC IMPLICATIONS: The findings of this analysis argue against the concern that poor adherence in Africa would be a rationale for delaying the expansion of antiretroviral therapy programmes. However, this does not mean that evaluation of adherence, and interventions to improve and maintain high levels of adherence, should not be instituted in African settings. Both are still required. Ensuring reliable drug supply and distribution networks are also important in maintaining high antiretroviral therapy adherence rates.

Source: USG/PEPFAR literature digests; <http://hivinsite.ucsf.edu/InSite?page=jl-00-00>, accessed on 31 January 2009.

APPENDIX FIVE: EXAMPLE OF AN AGENDA FOR AN EVALUATION PRIORITY-SETTING WORKSHOP

DAY 1

Time	Topic	Presenter/ facilitator
8:30–9:00	Welcome remarks Importance of a coordinated national evaluation agenda	Director of the national AIDS commission
9:00–9:30	Introduction: Workshop goals Review of the workshop agenda Workshop process and outputs	Facilitator
9:30–10:15	Overview of monitoring and evaluation data and country context: HIV epidemic National strategic plan for AIDS Questions & Answers	National AIDS commission representative & consultants
10:15–10:30	Tea break	
10:30–12:00	Overview of monitoring and evaluation data and country context (continued): Findings from available monitoring and evaluation data Evaluation studies to date Discussion	National AIDS commission representative & consultants
12:00–1:00	Lunch break	
1:00–3:30	Identify evaluation questions: Describe task and criteria for identifying evaluation questions Small group work by programme area to identify evaluation questions	Facilitator Small group facilitators

3:30–3:45	Tea break	
3:45–5:00	Report back from small groups: List evaluation questions by programme area and describe how they measure against the criteria used	Facilitator

DAY 2

Time	Topic	Presenter/ facilitator
8:30–10:15	Prioritize evaluation questions: Plenary review, classification and prioritization of top 10 evaluation questions from the small group work	Facilitator
10:15–10:30	Tea break	
10:30–12:00	Describe prioritized evaluation questions: Small group work to describe (a) relevance; (b) evidence; (c) actionability and programme utility for the prioritized evaluation questions	Small group facilitators
12:00–1:00	Lunch break	
1:00–2:30	Report back from small group work: Present the descriptions for each of the evaluation questions	Facilitator
3:30–3:45	Tea break	
3:45–5:00	Plenary discussion of small group work and agreement on final recommendations for evaluation priorities Discussion and agreement on next steps	Facilitator

NOTES

Publications in this series

Basic Terminology and Frameworks
for Monitoring and Evaluation
(2010)

An Introduction to Indicators
(2010)

A National Evaluation Agenda for HIV
(2010)

An Introduction to Triangulation
(2010)

ABOUT THE MONITORING AND EVALUATION FUNDAMENTALS SERIES

With the advent of the global financial crisis affecting most countries around the world, monitoring and evaluation (M&E) has become more important than ever before. Determining what programs do or do not work; implementing programs with proven cost-effectiveness; monitoring progress towards achieving targets; and ensuring accountability are objectives which are especially important now in the HIV response, as well as in other health and development areas. Thus, it is increasingly important that M&E is better understood, communicated in simplified language, and conducted in a coordinated and sustainable manner that generates information that can be easily used. Further, it is essential that M&E addresses the needs of and involves all key stakeholders right from the start and that results are made publicly available and utilized strategically in policy-making, planning, and program improvement.

This series provides a common sense introduction to a range of M&E issues. It covers the fundamentals and their practical applications and includes techniques and tools for managing M&E of the HIV epidemic and response. Although the series uses HIV as its focus, the M&E fundamentals are also relevant to other areas of public health and development. As such, these books may also be useful in strengthening national M&E systems designed to track progress in other health and development goals, such as those outlined in the United Nations Millennium Development Goals (MDG).