

**10**  
January

# Guidance on Developing Terms of Reference for HIV Prevention Evaluation



**UNAIDS**  
JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS

UNHCR  
UNICEF  
WFP  
UNDP  
UNFPA

UNODC  
ILO  
UNESCO  
WHO  
WORLD BANK

**MERG**  
MONITORING & EVALUATION  
REFERENCE GROUP



# Guidance on Developing Terms of Reference for HIV Prevention Evaluation



THE WORLD BANK



# Acknowledgements

This guidance has been developed under the auspices of the UNAIDS Monitoring and Evaluation Reference Group. The following members of the Technical Working Group for Prevention Evaluation (TWG-PE) contributed to this document [in alphabetical order]: Savina Ammassari (UNAIDS); Gregory Ashefor (NACA, Nigeria); Rachel Baggaley (Christian AID); Réne Bonnel (World Bank); Michel Carael (consultant); Henry Damisoni (UNAIDS), Nicole Fraser (World Bank/GAMET consultant); Wayne Gill (UNAIDS); James Guwani (UNAIDS); Eva Kiwango (UNAIDS); Cedric Mahe (Global Fund); Masauso Nzima (UNAIDS); Kayode Ogungbemi (PAHO/WHO); Kola Oyediran (Measure Evaluation); Rosalía Rodríguez-García (TWG-PE Chair/World Bank); Hernán Rosenberg (PAHO/WHO); Otilia Scutelnicu (MOH, Moldova); Ilene Speizer (Measure Evaluation); Matthew Warner-Smith (UNAIDS); Annick Wouters (consultant); Serge Xueref (Global Fund).

This guidance was validated by members of a broader virtual consultative group of the TWG-PE [in alphabetical order]: Savina Ammassari; Carlos Arnaldo, Harry Beauvais, Stella Bivol, Réne Bonnel, Henry Damisoni, Oluwole Fajemisin, Suriadi Gunawan, James Guwani, Todd Koppenhaver, Rene Leyva, Tendayi Ndori Mharadze, Julian Modley, Lydia Mukashyaka, David Mundy, Adjaratou Ndiaye, Stanis Ngarukiye, John Novak, Elizabeth Owiti, Kola Oyediran, Laura Porter, Hernán Rosenberg, Judith Shongwe, Ngak Song, Bob Verbruggen, Annick Wouters, and Barbara de Zalduondo.

Comments were provided during the piloting of this guide in Nigeria in September 2008 by the following participants [in alphabetical order]: Azeez Aderemi, Adeyinka Aderinto, Kayode Agada, Laura Arnston, Greg Ashefor, Mark Anthony, Lawal Bimpe, Okoeki Faith, Ivande Igundunasse, Victoria Isiramen, Isa Iyortim, Toyin Jagha, Orepekan Kemi, Abubakar Kende, A. Lawanson, Agbahoungba Lazare, Mukhtar Muhammad, Jummai Musa, Stanley Musah, Waterfield Ndak, Tessy Ochu, Victor Olisah, Mobolaji Olokodana, Offiah Olokodana, Kola Oloye, Lheadi Onwinkwe, David Onime Iwueze Rose, T. Sheikh, and Dooshima Uganden.

The TWG-PE is grateful to the following for their comments [in alphabetical order]: Paul Bouey (PEPFAR/OGAC), John Elder (World Bank), Jody Zall Kusek (World Bank), John Novak, (USG/USAID), Greet Peersman (Tulane University), Deborah Rugg (UNAIDS), Jack Reynolds (USAID consultant), Ray Rist (IDEAS).

# Contents

Acronyms	05
Purpose of the <i>Guidance</i>	06
Content of the <i>Guidance</i>	06
Intended users of the <i>Guidance</i>	07
<b>Introduction</b>	<b>08</b>
<b>1. The Importance of Rigorous Evaluation</b>	<b>09</b>
The Results Chain — A Framework for Evaluating HIV Interventions	11
Types of Evaluation	12
<b>2. Terms of Reference for Prevention Evaluation</b>	<b>14</b>
The Importance of a Systematic Approach to the Evaluation TOR	14
Planning the Evaluation	15
Prevention Evaluation TOR Template	16
How to Write the Different Sections of the TOR	16
1. Evaluation Topic: Evaluate the Behaviour Change Communication / HIV Counseling and Testing (BCC/HCT) Intervention	18
2. Background and Rationale: National AIDS Strategy Evaluation Requirements	19
3. Evaluation Purpose: Determine the Effectiveness of Prevention Interventions in Reducing Risk Behaviours	19
4. Users: Identify The Principal Users And Uses Of The Evaluation	20
5. Key Evaluation Questions: Use the Results Chain to Identify and Select Appropriate Evaluation Questions for the Relevant Population Groups	20
6. Target Group(S): Most-At-Risk-Populations and other Vulnerable Groups	22
7. Prevention Interventions: Describe the Specific Intervention(S) to be Evaluated	22
8. Prevention Indicators: Select at Least One Indicator for the Programme as a Whole or for Each Intervention to be Evaluated	23
9. Evaluation Design: Experimental, Quasi-Experimental, Non-Experimental	23
10. Key Data Sources And Procedures: Biennial Behavioural Survey and Clinic Records	24
11. Key Data Analysis Procedures: Performance Assessments, National and District Levels	26
12. Evaluation Activities and Schedule: Activities, Duration, Date of Final Report	26
13. Team Members and Number of Days: Internal or External Evaluation Team, Composition of the Team, Number of Days	27
14. Administrative & Logistical Support: Transport, Accommodation, Meeting Space, Other Support	28
15. Evaluation Budget: Total Amount and Budget Items	29
<b>Conclusion</b>	<b>31</b>

<b>Appendix 1: Illustrative Evaluation TOR Using the Template</b>	<b>32</b>
Sample: Fully developed TOR	32
1. Evaluation Topic	32
2. Background	32
3. Evaluation purpose	33
4. Users	33
5. Key evaluation questions	33
6. Target groups	34
7. Prevention interventions	34
8. Prevention indicators	35
9–11. Methodology: Evaluation design, key data sources, and analysis procedures	35
12. Evaluation activities and schedules	36
13. Evaluation team members and level of effort (LOE)	36
14. Administrative & logistical support	38
15. Evaluation budget (cost-sensitive material — do not divulge)	38
<b>Appendix 2: Illustrative Summaries of Evaluation TOR</b>	<b>39</b>
Illustration 1: Prevention outcome and impact evaluation	39
Illustration 2: Prevention outcome evaluation	41
<b>Appendix 3: Supporting Information</b>	<b>44</b>
A. EVALUATION APPROACH	44
Elements of Good Evaluation	44
B. EVALUATION DESIGNS	46
C. INTERNATIONAL GUIDING PRINCIPLES FOR EVALUATORS	51
Guidelines for Evaluation in Africa: The African Evaluation Association	53
D. COSTS AND COST-EFFECTIVENESS OF PREVENTION INTERVENTIONS	54
E. GLOSSARY OF M&E TERMINOLOGY	55
<b>References</b>	<b>57</b>

# Acronyms

AIDS	Acquired immunodeficiency syndrome
BCC	Behavioral change communication
CCM	Country Coordinating Mechanism (Global Fund)
CDC	US Centers for Disease Control and Prevention
FSW	Female sex workers
HCT	HIV Counseling and Testing
HIV	Human immunodeficiency virus
IDU	Intravenous drug user
IEC	Information, education and communication
LOE	Level of effort
M&E	Monitoring and evaluation
MDGs	Millennium Development Goals
MERG	Monitoring and Evaluation Reference Group
MOH	Ministry of Health
MSM	Men who have sex with men
NGO	Non-governmental organization
PLHIV	People living with HIV
PMTCT	Prevention of mother-to-child transmission of HIV
STI	Sexually transmitted infections
SW	Sex Worker
TOR	Terms of reference
TWG	Technical working group
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNDP	United Nations Development Programme
UNGASS	United Nations General Assembly Special Session on HIV/AIDS
UNICEF	United Nations Children's Fund
WB	World Bank
WHO	World Health Organization

## Purpose of the Guidance

This guidance was developed in response to the need for a simple tools and guidelines that countries can use to plan and implement rigorous evaluations of AIDS programmes. The specific objective is to foster a systematic approach to the evaluation of prevention programmes by focusing on an often overlooked, yet critical step, in evaluation planning: the preparation of terms of reference (TOR).

Note: some institutions use the term scope of work (SOW) rather than terms of reference; for all practical purposes, they are the synonymous.

This is not a step-by-step operational guide on how to conduct evaluations. The purpose of this guidance is to facilitate planning of evaluations of HIV prevention, discussions about the design of evaluations, and drafting of TOR to guide such evaluations. It assumes that the reader is familiar with the basics of HIV prevention and with the fundamentals of monitoring and evaluation to follow each of the TOR template steps.

The TOR provide an outline or template for the prevention evaluation. The outline is easy to use, adaptable, and not specific to a particular programme. It enables a clear definition of what needs to be evaluated, how and why; and, facilitates the selection of the evaluation team.

## Content of the Guidance

The TOR approach in this guidance is accompanied by brief instructions and supplemented by more detailed examples and background in the appendices. This document incorporates the comments gathered during field-testing with a representative group of the intended audience in Nigeria in September 2008. It also draws on the ideas and experiences shared during the 3rd UNAIDS Global Monitoring and Evaluation (M&E) Training in Bangkok in October 2008.

**Section 1** includes basic definitions and concepts about evaluation and explains the rationale for and role of evaluations.

**Section 2** describes how to prepare the terms of reference (TOR) for a prevention evaluation.

**Appendix 1** gives examples of evaluation TOR.

**Appendix 2** includes supporting information about evaluation including: evaluation standards and designs; international guiding principles for evaluators; cost and cost-effectiveness studies; and a glossary of commonly used M&E terms.



## Intended users of the *Guidance*

This guidance is intended for use by anyone who prepares or reviews TOR for evaluations of HIV prevention programmes. This includes: national AIDS programme managers, HIV prevention specialists, M&E advisors who are responsible for ensuring that evaluations are undertaken; programme and evaluation managers of nongovernmental organisations; staff of international agencies.

# Introduction

A comprehensive national AIDS programme includes prevention, treatment, care and support services. Treatment access has expanded steadily in recent years, but efforts to prevent new infections have lagged. As UNAIDS notes, the HIV epidemic cannot be reversed without strong sustained success in preventing new HIV infections as for every two people put on treatment, five people become newly infected with HIV (UNAIDS, 2008). Prevention remains the mainstay of the HIV response. But, the reality is that while 87 percent of countries with universal access targets have established goals for HIV treatment, only 50 percent of these countries have targets for key HIV prevention strategies (UNAIDS, 2008).

Success in HIV prevention influences the future trend of the epidemic, helps sustain treatment access, and mitigates the future cost to society resulting from new HIV infections (World Bank, 2002). National HIV prevention programmes need to include a mix of behavioural, biomedical and structural interventions appropriate for the characteristics of the HIV epidemic and the population groups that are most affected or most at risk. **Table 1** lists the 12 essential programmatic actions for HIV prevention recommended by UNAIDS.

Evaluation provides critical evidence on whether and how national HIV prevention programmes achieve results. It enables programmes to document how changes in behaviours and improvements in availability, access, utilization, and coverage of services result in decreased rates of HIV infection. It provides evidence for a critical base on which an effective HIV response can be built and scaled up to reach the goal of universal access to comprehensive prevention, treatment, care, and support (UNAIDS, 2007).

**Table 1: Essential programmatic actions for HIV prevention**

1. Prevent the sexual transmission of HIV
2. Prevent mother-to-child transmission of HIV
3. Prevent the transmission of HIV through injecting drug use, including harm reduction measures
4. Ensure the safety of the blood supply
5. Prevent HIV transmission in health care settings
6. Promote greater access to voluntary HIV counseling and testing while promoting principles of confidentiality and consent
7. Integrate HIV prevention into AIDS treatment services
8. Focus on HIV prevention among young people
9. Provide HIV-related information and education to enable individuals to protect themselves from infection
10. Confront and mitigate HIV-related stigma and discrimination
11. Prepare for access and use of vaccines and microbicides
12. Promote greater access to male circumcision

Source: UNAIDS (2005). *Intensifying HIV Prevention: A UNAIDS Policy Position Paper*. Geneva: UNAIDS, p.32. [Action 12 was added after publication of the position paper as an important programmatic action given recent evidence.]

# 1. The Importance of Rigorous Evaluation

In the development of terms of reference (TOR), the rationale for an evaluation is clarified and fundamental decisions are made about the evaluation questions and appropriate evaluation designs, approach, and implementation modalities. Developing the TOR, covered in detail in Section II, is a critical first step toward a credible evaluation. First, however, it is essential to understand the importance of evaluation and its place in programme planning.

An evaluation is conducted to determine the merit and worth of a policy, programme, or intervention through systematic collection of information about programme activities, characteristics, and outcomes. Evaluation studies provide credible information for use in improving programmes, identifying lessons learned, and informing decisions about future resource allocation.

Evaluations should be designed, whenever possible, as programmes are planned. This offers numerous advantages, including the opportunity to define realistic programme objectives, ensure complementarities between monitoring and evaluation in data gathering, determine baseline data, and ensure that activities needed to facilitate the evaluation are done in a timely manner.

Evaluations are conducted according to detailed evaluation protocols. The term protocol is used in medicine to refer to a scientific or medical experiment, treatment, or procedure. The term protocol in the social sciences refers to the detailed plan of a procedure—an evaluation in this case—that is based on accepted standards of practice.

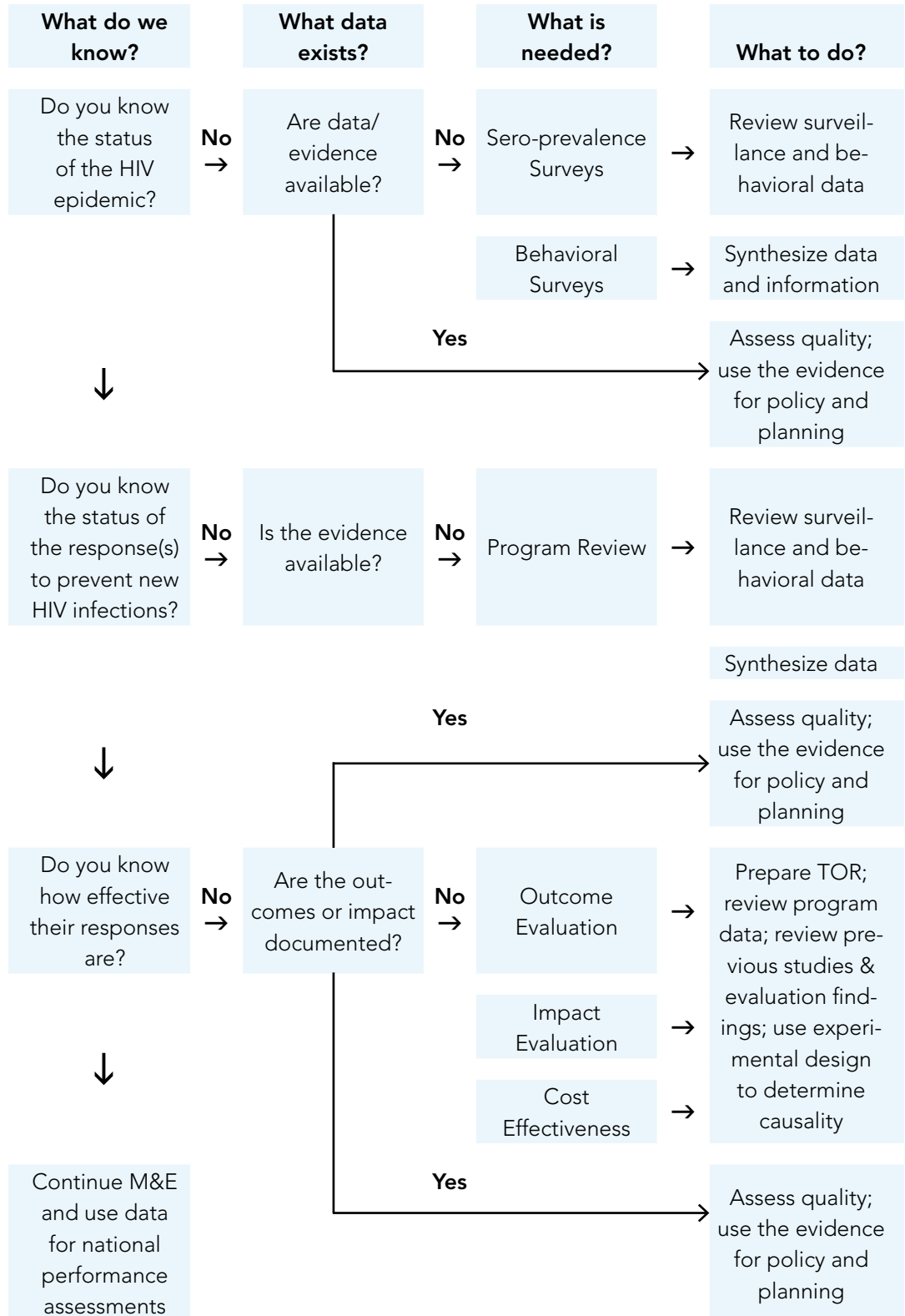
An evaluation protocol describes every step of the evaluation in detail, with special focus on the design, methodology, and analytical procedures of the evaluation. The protocol should include all sections of the terms of reference (TOR). In addition, it includes such items as (i) the make-up of the evaluation team; (ii) transfer of knowledge activities; (iii) the protection of human subjects if relevant to the evaluation; and (iv) an outline of the final evaluation report.

The evaluation protocol is not the proposal prepared in response to a bidding exercise, although the bidding proposal is likely to include components of the evaluation protocol. This is why the first technical product to be requested from the evaluation consultants is the evaluation protocol (sometimes also referred to as the *inception report*).

HIV prevention evaluation helps answer fundamental questions about trends in the national or local HIV epidemic and the effectiveness of the response. The decision tree in Figure 1 shows the key questions and actions that help build evidence for decision making.

**Figure 1.** Steps for identifying areas where strategic information is needed

Figure 1. Steps for identifying areas where strategic information is needed



When evidence does exist, it is important to assess its quality. Factors of quality may include the assumptions made in the studies, the analytical tools used, whether baselines or benchmarks existed and were applied, and the validity of the evidence — that is, the extent to which the data collection strategies and instruments measure what they purport to measure.

Increasing the emphasis given to HIV prevention evaluation is important for several additional reasons:

- There is a need to analyze and document which prevention interventions are effective and should be scaled up to prevent new HIV infections. In recent years there have been more improvements in monitoring than in evaluation, and the imbalance is continuing to grow (Adamchak et al., 2004).
- More knowledge is needed on most-at-risk populations (i.e., sex workers, injecting drug users, men who have sex with men). Many HIV prevention interventions are aimed at these groups, especially in concentrated epidemics, but very few evaluations distinguish among these groups according to the risk of transmission.
- There is a need to know more about the relative costs of prevention interventions and their cost-effectiveness (e.g., the cost per infection averted). This information is crucial to the allocation of resources and strategies to scale up certain interventions.

## The results chain — A framework for evaluating HIV interventions

A results chain —also called a logic model or logical framework— is used to clarify assumptions about how activities contribute to achieving results. Results chains are results-oriented, demonstrate change, and show logical connections. The results chain is based on the if-then theory or logical progression that if X is done or happens, then Y will follow as the logical consequence. For example, if pregnant women are tested and counseled and HIV-positive women are treated, then mother-to-child transmission of HIV will be reduced.

**Figure 2** summarizes such a logical chain of results, in which resources (inputs) are processed into goods and services (outputs). These result in knowledge and behavior changes and improvements in access to and utilization of services (outcomes), which, in turn, eventually produce changes in the socio-demographic or epidemiological profile of a population (impact).<sup>1</sup>

Each step of the results chain must be measured. If results are not achieved or changes are not observed in the first steps, then there is no need to look for or expect changes further down the line. Each step of the results chain must be measured in order to determine the plausibility of observed outcomes being attributable to the intervention, especially in the absence of control groups.

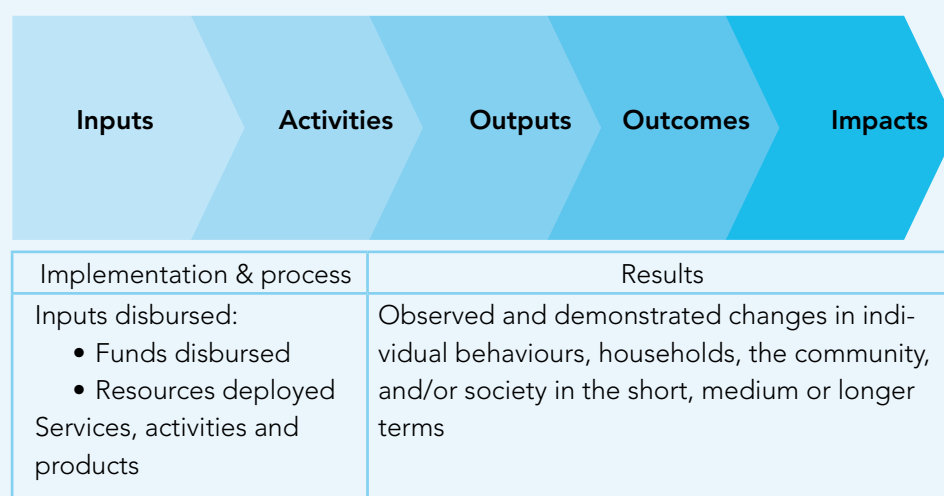
---

1 This framework goes back at least to 1975 and is still relevant and used. See Jack Reynolds, et. al. Evaluation Handbook for Family Planning Programs, U.S. Department of Health, Education and Welfare, 1975; UNAIDS. A Framework for Monitoring and Evaluating HIV Prevention Programmes for Most-At-Risk Populations, April 2007, p. 5; World Bank. An Evaluation of Bank Support for Decentralization in Client Countries, Fast Track Brief, 2008; and World Bank GHAP. Health Systems and HIV/AIDS: Impact on Service Delivery (working draft), 2008.

The outputs of a programme or intervention are largely controlled by management through the allocation and use of resources, the provision of services, the organisation of activities, and the products completed through the implementation processes. These processes are logically expected to lead to the achievement of results or changes —outcomes and impacts— though the latter are typically the result of more than one programme.

An outcome evaluation is 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/comparison groups are not available, e.g., for the evaluation of a national programme.)

Figure 2: Framework for evaluation: The causal links of the results chain



Developing an evaluation framework using the results chain will help clarify the goals and expected results of the programme or intervention to be evaluated. It will also help clarify whether the intervention has been implemented in accordance with international standards.<sup>2</sup> To assess how well a programme is doing, the TOR should be clear in specifying which of the links in the chain will be assessed.

## Types of Evaluation

This guidance focuses on four types of evaluation: (1) programme logic model evaluations; (2) process evaluations; (3) outcome evaluations; and, (4) impact evaluations.

<sup>2</sup> UNAIDS is in the process of defining each prevention intervention to systematize how each intervention is conceptualized, implemented and reported on. This is crucial for evaluation in general as well as for cross-national analyses.

- **Programme logic model (or result chain) evaluations** are used to determine the logic of the causal model behind the HIV prevention intervention under study. The evaluation addresses each step of the results chain, starting with the amount and type of inputs and the sequencing of activities that are used to bring about a desired change. The evaluation focuses on the plausibility of achieving the desired change based on previous field experience and published evidence (Kusek and Rist, 2004). Existing evidence and theory of change are used to make sure that the design of the intervention has potential for success.
- **Process evaluations** focus on program implementation, including how services are delivered, differences between the intended population and the actual population served, access to the programme, service quality assessments, and management practices. In addition, process evaluation might provide understanding about the cultural, socio-political, legal, and economic contexts that affect programme implementation.
- **Outcome evaluations** are concerned with determining if, and by how much, programme activities or services have achieved their intended outcomes among the targeted population. For example, does the percentage of factory workers using condoms increase after a workplace project has made them aware of how HIV is transmitted? And how large is the increase? Have men who have sex with men changed their sexual behaviours because of a peer counseling intervention? Has the percentage of injecting drug users using sterile needles increased due to a needle exchange program?
- **Impact evaluations** are a scientifically rigorous methodology to establish a distal causal association between programmes and what they aimed to achieve in terms of disease reduction. The long term effects (impacts) seldom can be attributed to a single programme or even a few programmes. Evaluations of impact on populations usually entail an evaluation design that includes the combined effects of a number of programmes and modeling of the long term and projected effects.

Field work shows that the term “impact evaluation” is often used more loosely than the restricted scientifically rigorous evaluation defined here.

All types of evaluation are important, and the TOR template described in this guidebook can be used to plan any type of evaluation. However, the intention of this guidance is to promote evaluations that build evidence of results in terms of outcomes and impacts. Evaluation of key processes and activities should always be included to determine whether interventions were carried out as planned.

The TOR are an important part of this process, and they can also be used for the preparation of cost-effectiveness studies. Cost-effectiveness studies analyze tangible benefits produced by money spent (see **Appendix 1**).

## 2. Terms of Reference for Prevention Evaluation

This section describes how to prepare terms of reference (TOR) for a prevention evaluation. To prepare good evaluation TOR, it is useful to have some ready-made templates or checklists that managers, prevention specialists, and M&E advisers can use. A template needs to cover stakeholder expectations for the evaluation, the purpose of the evaluation, and key questions to be answered. It also needs to include the HIV prevention interventions, data needs, indicators, staffing requirements, and the evaluation schedule and budget. The TOR template provided here can be adapted to any prevention evaluation (see **Appendix 1**, Illustrative summaries of evaluation TOR).

### The importance of a systematic approach to the evaluation TOR<sup>3</sup>

Whether the evaluation is implemented in-house or is outsourced, the TOR provide a systematic approach to planning the evaluation. They are a catalyst for more in-depth and concrete dialogue among evaluation stakeholders and facilitate the preparation of a responsive evaluation protocol because expectations are made explicit and are documented.

Although there are a number of generic guidelines and checklists<sup>4</sup> for preparing TOR for evaluation, none focus on *how to prepare TOR for prevention evaluation* nor *who* should prepare them. The result is, that TOR are often put together in an ad hoc way. They may lack key elements, such as what data should be collected and how these data should be analyzed, or have unrealistically long lists of evaluation questions to be answered in unrealistic timeframes.

What is needed is a standardized approach to developing TOR for HIV prevention evaluations, so that those who prepare them and those who review them, use the same comprehensive, but not excessive, reference base. In the pilot of this guidance and TOR template by approximately 35 HIV prevention and M&E technical staff and managers from the public sector, nongovernmental organizations (NGOs), academia, and donor agencies in Nigeria, participants identified a number of reasons to have a standardized approach to planning prevention evaluations and developing the TOR.<sup>5</sup> Reasons included the following:

- The steps in the template stimulate debate and discussion, which help to improve the content and quality of the TOR and ultimately the evaluation.
- The TOR template helps the evaluation team to reach agreement on key terms, constructs, and perspectives.
- The preparation of TOR should not be left to M&E specialists alone but should be done by a stakeholder team that also includes HIV prevention specialists, management, and finance personnel. By involving everyone in the process, the whole organisation reaches a common understanding of what it means to plan and conduct an evaluation. Moreover, involving stakeholders in the planning of an evaluation enhanc-

<sup>3</sup> TOR are prepared for a variety of reasons and purposes. In this guidance the focus is on TOR for HIV prevention evaluations, hence the focus is on being systematic and rigorous as they describe the framework that frames and guides the evaluation. Note that some institutions use the term “scope of work” (SOW) instead of TOR.

<sup>4</sup> For example: Checklists for Program Evaluation Planning by McNamara; Evaluation Contracts Checklist by Stufflebeam; and Key Evaluation Checklist by Scriven.

<sup>5</sup> The pilot was conducted in September 2008 in Nigeria.



es the utilization of the evaluation results afterwards as the intended users will have a stake in the findings (Quinn-Patton, 1985).

- Management and finance personnel gain a better understanding of what is involved in planning an evaluation and what it may require in funds, personnel, and field work.
- Prevention specialists and M&E advisers need to work together to better understand the programme and interventions that are to be evaluated.
- The utilization by any organisation (public, private, nongovernmental or civil society, donor agency) of the same TOR template for planning different types of evaluations would add rigour and systematize knowledge building efforts.

## Planning the evaluation

The TOR are a critical first step towards implementing robust evaluations. Once the TOR are agreed by stakeholders, the national AIDS programme or the concerned organisation will follow the TOR with the preparation of a detailed evaluation protocol. If the intention is to outsource the evaluation, then the TOR provide the basic information for the preparation of the call for letters-of-interest and/or bids. If the intention is for the evaluation to be conducted by an internal team, then this team will need to develop the TOR further into the detailed evaluation protocol. **Figure 3** below provides an illustration of the logical steps in the process from the preparation of the TOR to the completion of the evaluation protocol, the implementation of the evaluation and the dissemination of the findings. The actual time needed, will vary in each situation depending on the complexity of the evaluation.

**Figure 3: Illustrative key steps from evaluation planning to dissemination of findings**

TOR Stage	Contracting Stage	Implementation Stage
1. Consult evaluation agenda on HIV or M&E plan	1. If outsourced, request bids to respond to TOR	1. Begin evaluation
2. Launch evaluation planning with stakeholders	2. Review bids and select consultant(s)	2. Field test of tools
3. Identify the person(s) to draft TOR	3. Request comprehensive evaluation protocol from consultants as product # 1 of their contract	3. Revise tools
4. Develop evaluation TOR	4. Develop evaluation protocol	4. Conduct field work, data coding, cleaning
5. Review & revise TOR	5. Review protocol and feasibility of implementation, including data collection tools & revise	5. Analyze findings and determine programme implications
6. TOR agreed with stakeholders	6. Evaluation protocol agreed with stakeholders	6. Prepare findings for validation with stakeholders
		7. Complete draft evaluation report for review
		8. Finalize report
		9. Hold evaluation dissemination seminar

## Prevention evaluation TOR template

**Table 2** is an example of a generic template for the terms of reference (TOR) of an evaluation. The focus is on documenting results achieved and changes observed, with the emphasis on outcomes. However, the template can be used for the preparation of TOR for a variety of evaluations.

**Table 2: Prevention evaluation TOR template**

1. Evaluation topic
2. Background and rationale
3. Evaluation purpose
4. Users of the evaluation
5. Key evaluation questions
6. Target group(s)
7. Prevention interventions
8. Prevention indicators
9. Evaluation design
10. Key data sources and procedures
11. Key data analysis procedures
12. Evaluation activities and schedule
13. Evaluation team members and level of effort (LOE)
14. Administrative & logistical support
15. Evaluation budget

The TOR need to include a number of key sections. Here they have been clustered into 15 sections, each of which should be concise, complete, and clear. However, sections may be collapsed, and a shorter TOR may be justified depending on the scope of the evaluation.

One way to start the process of developing the TOR is to prepare an annotated outline first, like the one shown in **Table 3** below — the TOR template summary — which summarizes each step in a brief narrative. This can be used for the first evaluation meeting of stakeholders to make sure that everyone is on the same page. Another important objective of the TOR template summary is to help stakeholders grasp the scope of the evaluation so that a realistic timetable can be developed at this time. It is advisable to identify the financial resources and other contributions available for the evaluation at this time. Once the outline is agreed upon, each section of the TOR can be completed in more detail.

## How to write the different sections of the TOR

The person responsible for making sure that the evaluation is done, will use the TOR template summary to clarify rationale, expectations, scope of work, time frame, and resources available and will form a team responsible for drafting the full TOR.

This evaluation planning team would include not only M&E personnel but managers, prevention specialists, and other key personnel from such areas as finance, data management, and logistics. It may include other stakeholders as appropriate. The team may also decide that some sections of the TOR would be further developed by the evaluators or the consultant(s) if outsourced. In most cases, it is expected that most of the sections will be written up by the team and that there will be a review of each section. In the following section, an illustrative example is used.

The overall TOR are summarized in **Table 3**. The example focuses on the evaluation of a behaviour change communication programme to increase HIV test-seeking behaviour and condom use.

<b>Table 3: Example of a prevention evaluation TOR template summary</b>	
<b>Evaluation topic</b>	Behavioral Change Communication / HIV Testing and Counseling intervention
<b>Background/ rationale</b>	National AIDS strategy requires biennial evaluations
<b>Evaluation purpose</b>	To determine effectiveness of the programme in increasing HIV test-seeking behaviour and condom use
<b>Users</b>	National AIDS Coordinating Authority, Ministry of Health, Country Coordinating Mechanism, UNAIDS, WHO
<b>Key evaluation questions</b>	How effective has the programme been in increasing the desired behaviour changes in the study populations?
<b>Target group(s)</b>	Female sex workers, their clients, men who have sex with men, people living with HIV
<b>Prevention interventions</b>	HIV counseling and testing; peer education; condom promotion
<b>Prevention indicators</b>	Percentage individuals knowing their HIV status (by age and by most-at-risk population); Percentage of condom use at last sex
<b>Evaluation design</b>	Quasi-experimental, lagged design by district
<b>Key data sources and procedures</b>	Biennial behavioural surveys, clinic records
<b>Key data analysis procedures</b>	Programme data assessments, national and district levels
<b>Evaluation activities and schedule</b>	Time will vary depending on the complexity of the evaluation from 45 to 90 days or more
<b>Team members and no. of days</b>	HIV prevention specialist, counseling and testing specialist, epidemiologist total = 60-90 days
<b>Administrative &amp; logistical support</b>	Organisation is responsible for local transport, hotel, meeting space; All other support is to be provided by the contractor.
<b>Budget</b>	USD amount plus in-kind contributions

The steps of the template are further described below.

## 1. Evaluation topic: Evaluate the Behaviour Change Communication / HIV Counseling and Testing (BCC/HCT) Intervention

The first step is to determine what is to be evaluated. In some cases the evaluation may have been scheduled when the national AIDS program was developed. In other cases, an ad hoc request may have been made for an evaluation of a particular programme, project, or intervention. Consult the national AIDS strategy and/or the national HIV prevention strategy, and/or the national M&E plan to determine if evaluations are already scheduled in. If an evaluation has not been scheduled, clarify the rationale for the evaluation.

### → DO:

- Identify the theory of change and logic model underlying the programme or intervention.
- Describe the programme or intervention to be evaluated.
- Decide on the scope of the evaluation in terms of its geographic area (e.g., national, provincial, state, district, community).
- Identify the intervention levels: program, project, intervention, activity.

### **Using the Results Chain, Draw or map the theory of change — or results chain or logic model — for the programme, project, or intervention that is to be evaluated.**

Mapping is the graphic representation of the elements of a program arranged to show the relationships among program inputs, activities, outputs, outcomes, and impacts. Mapping is a shorthand way to describe the logic or theory of change for the programme. This can be very helpful for the team that is preparing the TOR as well as for those who need to prepare a proposal to respond to the TOR or to prepare the actual design of the evaluation protocol. Mapping can be done in a variety of ways. Here the suggestion is to use the results chain to make the assumptions behind the intervention explicit so that they can be tested to learn over time whether they are valid or not. This example illustrates how the key elements of the BCC/HCT intervention can be identified for the evaluation (**Table 4**).

**Table 4: A simple results framework for a BCC/HCT intervention**

<b>Inputs →</b>	<b>Activities →</b>	<b>Outputs →</b>	<b>Outcomes →</b>	<b>Impact</b>
Funds provided	Contacting target groups	Target groups reached	Risk reduction  Increased correct condom use	HIV infections reduced
Brochure acquired	Distributing brochures	Number of brochures distributed		
HIV test kits purchased	Distributing HIV test kits	Number of HIV tests conducted with counseling		
Condoms acquired	Distributing condoms	Number of condoms distributed		

## 2. Background and rationale: *National AIDS strategy evaluation requirements*

This should include a background statement on the context for the evaluation. At a minimum, it would cover the following:

- A brief overview of the HIV epidemic in the country, including prevalence, incidence, and trends. Possible sources of data may include: national indicators, surveillance/survey data, UN General Assembly Special Session on HIV/AIDS (UNGASS) reports, national AIDS strategy, HIV prevention strategy, situational analysis done for a Global Fund proposals.
- A description of the HIV response including government, nongovernment, and donor contributions.
- Achievements to date and plans for the future.
- Other evaluations done to date and what can be learned from them.
- Need for the evaluation and additional evidence for HIV prevention.

## 3. Evaluation purpose: *Determine the effectiveness of prevention interventions in reducing risk behaviours*

State the specific outcomes (and impacts, where appropriate) to be assessed. In most cases this should conform to the kind of evaluation specified. For example, for impact evaluations you may select a health objective (e.g., HIV infection); for outcome evaluations you may select a behavioural objective (e.g., correct and consistent condom use). The geographic area and the time frame for the evaluation need to be defined as well. Some examples of evaluation purposes are:

- To determine the effectiveness of HIV prevention interventions in reducing risk behaviours among most-at-risk population groups in the targeted areas.
- To determine if there has been a significant reduction in co-infections (HIV/TB) in the past year among female sex workers enrolled in sexually transmitted infection clinics throughout the country.

It is also important at this stage to clarify broadly how the evaluation findings will be used. For example, are the evaluation findings intended to help policy makers decide whether a programme or particular intervention for a specific target group should be continued or phased out or scaled up or down, and/or be replicated in other settings?

#### 4. Users:

##### *Identify the principal users and uses of the evaluation*

An important element of any evaluation that needs to be addressed early on in the planning stage is the identification of different users (audiences) of the evaluation findings and how they intend to use them. It is essential to involve these stakeholders periodically throughout the evaluation planning and implementation process, as well as at the end of the evaluation when the results are in.

→ **ASK:** Who is going to use this evaluation? Who are the key decision makers? List the principal audience(s) at the local, national, and international levels. Include private as well as public organisations. For example: the national AIDS coordinating authority, programme managers, the Ministry of Health, M&E advisers, the country coordinating mechanism, local nongovernmental organizations, the private sector, development partners.

#### 5. Key evaluation questions:

##### *Use the results chain to identify and select appropriate evaluation questions for the relevant population groups*

The key questions for evaluation are dictated by the topic of the evaluation. At this stage, the results chain needs to be applied to determine whether implementation has occurred as planned and to identify the key questions and assumptions that the evaluation should address. Start by clarifying whether the evaluation is related to a programme, a project, or an intervention, and what population groups are targeted. Then, define the key results expected of the programme as indicated in the results chain and select the questions that are relevant for assessing whether or not the key results have been achieved. Monitoring data must be consulted to establish what is already known. Subsequently, identify the appropriate sequence of the evaluation questions. A possible question may be: *How effective has the programme been in reducing risk behaviours in sex workers?* There are likely to be some sub-questions about the effects disaggregated by specific settings or by age,

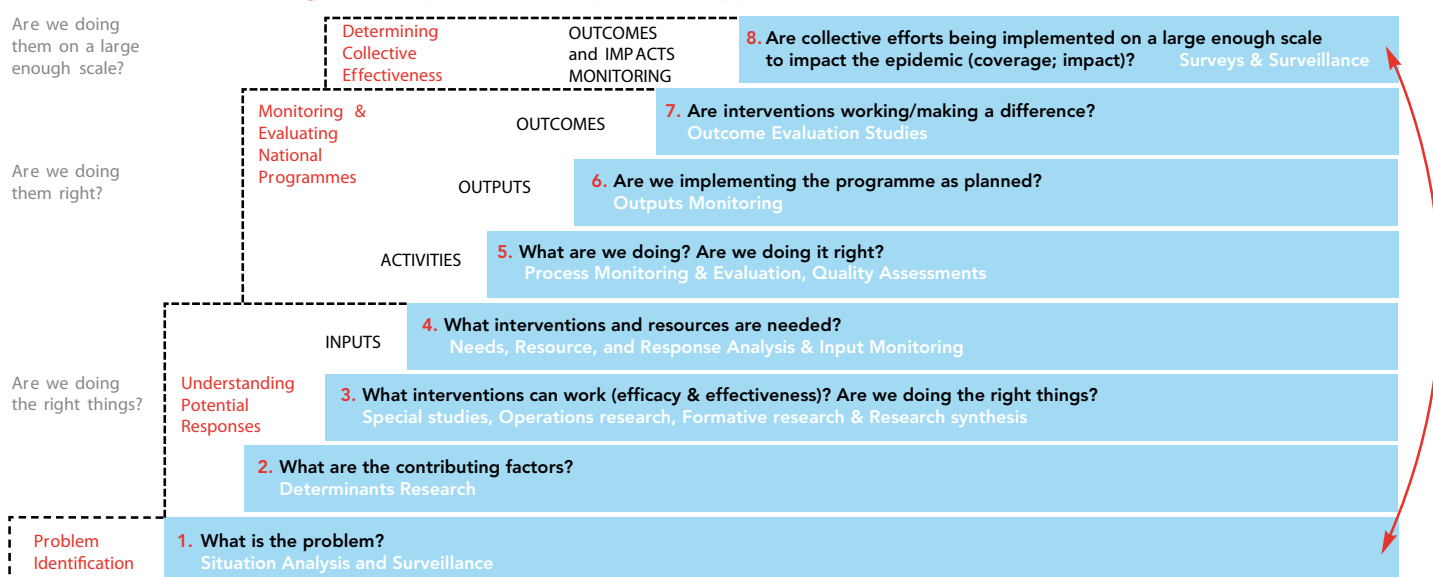
education, etc. These can be called “**Basic or Level 1**” questions. These may be all the decision makers want to know about. If so, resist the temptation to add more questions, as data needed to answer each additional question may have cost and time implications.

“**Level 2**” questions try to determine why behaviours and/or new HIV infections did or did not change from the baseline level. *What were the key factors, within the programme and beyond, that affected this observation? Were the inputs obtained and the activities implemented as planned?; Were the expected outputs achieved?* Note that these are all process questions, not outcome or impact questions. They are crucial to understand how and why results were/were not achieved.

“**Level 3**” questions look towards future trends and/or whether an intervention can be expanded or replicated. Questions of programme costs and cost-effectiveness or funding flows may also be appropriate to determine whether scaling up of the particular programme is advisable.

Throughout this process of planning, reflection, and decision-making, those developing the TOR of the evaluation need to consult national documents (i.e., the national AIDS strategy, the national HIV prevention strategy, the national M&E plan), monitoring data (i.e., facility-based and community-based), previous evaluations (by local and international organisations), surveys and special studies (i.e., Demographic and Health Survey, Modes of Transmission study, epidemiological synthesis), HIV-related reports, and whatever reliable information already exists. These will elucidate the topic under consideration and help the stakeholders decide whether the evaluation is indeed warranted, and if so, what the fundamental evaluation questions and scope are. **Figure 4** shows a hierarchy of evaluation questions based on the results chain.

**Figure 4: A public health questions approach to HIV M&E**



Source: Rugg, D., Peersman, G., & Carael, M. (Eds.). (2004). Global advances in HIV/AIDS monitoring and evaluation. *New Directions for Evaluation*, 2004 (103).

## 6. Target group(s):

### *Most-at-risk-populations and other vulnerable groups*

In planning the evaluation, be specific in identifying the target group(s) of the prevention programme. For example, prevention programmes may address sex workers (SW), men who have sex with men (MSM), injecting drug users (IDU), people living with HIV (PLHIV), high-risk youth, and migrant workers. But some programmes may target all adolescents or the general population, etc. It is important to specify the gender, age, cultural, and other characteristics of the target groups. Also identify their geographical location, as this will have an impact on the costs on the evaluation.

It may be useful to distinguish HIV prevention efforts targeted to populations who are not infected with HIV, where the goal is to avoid primary infections (i.e., prevention of initial HIV infection), and efforts targeted at HIV-positive people to prevent secondary infections (i.e., prevention of transmission of HIV from the infected person to another person).

For secondary prevention (sometimes referred to as “positive prevention”), it is especially important to reach people who have the potential to transmit HIV to large numbers of other people (i.e., people living with HIV who have multiple sex partners or share needles with other people). The evaluators should use existing evidence to be as specific a possible about the target population(s) in evaluation questions.

## 7. Prevention interventions:

### *Describe the specific intervention(s) to be evaluated*

An essential question needs to be answered: Is this an evaluation of the overall programme or of one or more components of the programme? If the former, state so in the template. If the latter, decide which programme components to include and list them; also indicate whether these components will be assessed together or individually. (For key interventions, see **Table 1**.)

As previously mentioned, it is important to apply the results chain to map out the logic of the implementation of the programme that is being evaluated. Here, it is advisable to consult UNAIDS guidelines on HIV prevention<sup>6</sup> and the national HIV prevention strategy to ascertain that the recommended benchmarks have been respected.

---

<sup>6</sup> You may wish to consult: UNAIDS Policy Position Paper for Intensifying HIV Prevention, 2005; UNAIDS Practical Guidelines for Intensifying HIV Prevention-Towards Universal Access., 1997; and A Framework for Classifying Prevention Interventions, 2008.



## 8. Prevention indicators:

*Select at least one indicator for the programme as a whole or for each intervention to be evaluated*

When the evaluation design is formulated, you will need to select at least one indicator for the programme as a whole or for each component intervention to be evaluated. The more indicators are included, the higher the cost and the more time needed to collect, tabulate, and analyze the required data.

Here are some examples of indicators relating to outcomes and impacts:

### **Changes in behaviour: Outcome indicators**

- Number of sex partners (specify whether concurrent or total partners) in the past 12 months.
- Percentage of young men (specify age groups) who have ever had sex.
- Percentage of sex workers using condoms consistently.

### **Changes in health status: Impact indicators**

- HIV sero-prevalence among the general population.
- HIV sero-prevalence among injecting drug users.
- HIV sero-prevalence in pregnant women aged 15 to 24 years old.
- Number of new infections in specific populations.

For process evaluations you will also need indicators for inputs, activities, and outputs. For example:

### **Activities conducted: Output indicators**

- Number of pregnant women attending antenatal care who have been tested for HIV.
- Number of sex workers who receiving peer education.
- Number of condoms distributed.

## 9. Evaluation design:

*Experimental, quasi-experimental, non-experimental*

The choice of the evaluation design depends on the evaluation context and needs; the key questions should drive the choice of evaluation methods. Many evaluations are non-experimental (i.e., do not include a control or comparison group). Quasi-experimental and experimental designs are more complex and can be more expensive, and are not always warranted to answer the questions under study. However, when rigorous evaluations are warranted, outcome evaluations should attempt to identify appropriate comparison groups to show the effect of the programme. If an experimental or quasi-experimental design is warranted, you may want to consult an expert to help develop an appropriate design.

The three most common non-experimental designs are: (i) one-shot case studies; (ii) one group pretest–posttest; and, (iii) time series. Evaluation designs are usually presented with O (i.e., observation of the status of the target group) and X (i.e., intervention to improve status) with time passing left to right. Thus, O X O indicates a pretest observation, introduction of the intervention, and a posttest observation to measure any change after the intervention occurred (**Figure 5**).

**Figure 5: Common non-experimental designs**

Study design	Intervention (X) and Observation (O)				
One-shot case study		X	O		
One-group pretest-posttest	O	X	O		
Time series	OOO	X	O	X	OOO

Quasi-experimental designs can be diagrammed by adding what happens to a comparison group (see **Appendix 2**). Some interventions have already been subject to controlled experiments and have been shown to work in which case there is no need to evaluate the intervention again with a rigorous evaluation design.

## 10. Key data sources and procedures: *Biennial behavioural survey and clinic records*

There may actually be data already available (see **Table 5**) and it is, therefore, recommended to first identify the most relevant secondary data sources for understanding the context as well as for use in the evaluation.<sup>7</sup> These would normally include routine monitoring data, project or programme strategic plans, annual or quarterly reports, recent evaluations, recent surveys, and special reports. In some cases, there may be ongoing studies that could be useful for the evaluation. Donor agencies may also be a source of relevant data, especially data that has been collected for their own purposes.

**Table 5: Potential data sources**

<ul style="list-style-type: none"> <li>• Human development reports</li> <li>• UNGASS reports</li> <li>• MDG reports</li> <li>• WHO 3 by 5 reports</li> <li>• AIDS programme effort indices</li> <li>• Spectrum reports</li> <li>• Goals reports</li> </ul>	<ul style="list-style-type: none"> <li>• Demographic and health surveys</li> <li>• Behavioral surveillance surveys</li> <li>• National AIDS strategies</li> <li>• Proposals to donors</li> <li>• World Bank analytical papers</li> <li>• Peer reviewed articles</li> </ul>
--	--

<sup>7</sup> Data that have already been collected and reported are called “secondary data”. “Primary data” are those that have to be collected and processed as part of the evaluation being planned.

Most evaluations require the collection of new data. This involves identifying the types of data that need to be collected and the sources of these data first, followed by developing data collection procedures. The principal options are:

**Direct observations:**

- Unstructured (e.g., ad hoc site visits to observe peer education process)
- Structured (e.g., a checklist to observe interaction between staff and a client)

**Direct reports:**

- Unstructured interviews (e.g., ad hoc conversations)
- Semi-structured interviews (e.g., a list of key questions for group interviews/discussions, key informant interviews)
- Structured interviews (e.g., a pre-coded questionnaire)
- Tests (e.g., blood tests, certification tests)
- Inventories (e.g., condom supplies in store)
- Focus groups

**Records:**

- Statistical data (e.g., number of HIV counseling and testing sessions)
- Documents (e.g., evaluation reports)

Qualitative data sources are included above and should be given special attention. Often, site visits, unstructured interviews and observation etc. are very helpful for identifying the internal and external factors that affect programme achievements.

It may be useful to develop a matrix that lists the evaluation question, evaluation sub-questions, measures or indicators, baseline data, evaluation design, data sources, and data analysis (see **Table 6** for an example).

**Table 6: Illustrative evaluation design matrix**

Evaluation Question	Evaluation Subquestion	Measures or Indicators	Baseline Data?	Evaluation Design	Data Sources	Data Analysis
How effective has the programme been in reducing risk behaviours among MSM?	Has there been a reduction in concurrent sexual partners among MSM?	Number of concurrent sexual partners in the past year	Yes, add value and year	Quasi-experimental, time series	Programme and survey data	Trend analysis

## 11. Key data analysis procedures: *Performance assessments, national and district levels*

At this point in the development of the evaluation TOR, the indicators selected to measure the evaluation questions have driven the methodological design and the data collection plan. Now it is time to plan the type of data analyses needed. There are two main analytical procedures: quantitative and qualitative analysis. They can both be applied to primary as well as secondary data. Quantitative analysis uses statistics to allow for comparison and tests of significance of the causal relationships observed. Qualitative procedures can enrich a quantitative analysis by providing a contextual analysis that help stakeholders understand the “why” of human behaviour. Other analysis procedures are:

- Cost analysis (amount and percent of expenditures by category; through monitoring).
- Cost-effectiveness analysis (also known as efficiency), which compares results and costs of different approaches; the approach with the lowest cost per unit result is the most efficient.

Determining how cost-effective an intervention is requires an estimation of the total cost of the intervention. Unit costs are not too difficult to compute if the denominator is an output, such as the cost per pregnant woman tested, but they can be difficult to compute if the denominator is difficult to measure. For example, the cost of an infection averted or the cost per injecting drug user adopting risk-avoidance practices are very difficult to measure (see **Appendix 2**). However, this information would be extremely valuable for managers who want to maximize resources (i.e., support interventions with the greatest effects at the lowest cost).

## 12. Evaluation activities and schedule: *Activities, duration, date of final report*

After defining what is needed, the next step is to list the necessary evaluation activities and their corresponding time schedule. A timetable chart is a useful way to summarize this information. The chart should include, at a minimum, the activities and their start and end dates. It can be expanded to include personnel requirements and level of effort. **Table 7** provides an example of a simple timetable estimating the number of days work within a nine-month period.

**Table 7: Illustrative timetable of evaluation activities**

Activity	Months								
	1	2	3	4	5	6	7	8	9
Preparation	3 days								
Team planning		4 to 8 days							
Field work			10 to 15 days	10 to 15 days					
Data analysis & review					10 to 20 days				
Report preparation						8 to 12 days			
Debriefing; revisions							2 to 5 days	2 to 5 days	
Finalize report									5 days

### 13. Team members and number of days: *Internal or external evaluation team, composition of the team, number of days*

The evaluation team or manager decides whether the evaluation is to be done internally (i.e., by the organisation itself) or externally (put out for bids). In either case, the TOR would specify the skills requirements. If the work is to be done externally, bidders would identify and propose specific individuals. Building on the evaluation activities and schedule from the previous step, it can be useful to prepare another table that summarizes the type of expertise and number of work days needed for each activity. Part of this exercise will be to clarify roles and responsibilities of the staff involved in this evaluation. An example for an evaluation team is presented in **Table 8**.

Table 8: Illustrative summary of team members and number of days

Activity and number of days								
Expertise	Preparation	Team planning	Field work	Data analysis/ review	Report preparation	Debriefing & revisions	Finalize report	Total days / person
Evaluation manager	3	6	2	2	2	2	2	19
Team leader / HIV prevention specialist	3	6	20	4	10	2	5	50
HIV prevention specialist	3	6	10	4	10	2	1	28
Epidemiologist	1	6	10	10	10	2	2	40
Data collectors	1	3	20	5	6	2	1	28
Statistician		6	2	10	1	2	1	28
Budgeting specialist		3	1	1	5	2	5	10
Writer/Editor								12

#### 14. Administrative & logistical support: *Transport, accommodation, meeting space, other support*

Oversights in the logistical planning of an evaluation can create serious problems during implementation and field work. The TOR should spell out the responsibilities, not only of technical staff, but also of the administrative and logistical support staff of the sponsoring and co-sponsoring organisations. If the decision has been made that the evaluation will be outsourced, then the TOR need to include the major expectations from the bidders.

The details of the bidders' responsibilities will be spelled out in greater detail in the bidding documentation, not in the TOR for the actual evaluation. Bidding documentation will include information such as: who is responsible for getting data collection clearances, organizing travel and accommodation, booking meeting space, making appointments for interviews, etc. **Table 9** shows a sample checklist that can be helpful for describing who will be responsible for what support activity.

**Table 9: Illustrative logistics support checklist**

Function	Sponsoring organisation(s)	Contractor	Other (specify)
International transportation			X
Visas			X
Local transportation-Cars/drivers		X	
Insurance		X	
Hotel arrangements		X	
Business appointments	X		
Meeting space	X		
Audio-visual and IT support	X	X	
Office space and equipment		X	
Printing and photocopying		X	
Interpreters/translators		X	
Report production		X	
Report dissemination	X		

## 15. Evaluation budget: *Total amount and budget items*

The final step in the development of the TOR is the preparation of the budget. Here it is essential to be comprehensive including all cost items. The objective is to be realistic about the actual costs and not to under or over-budget. The key activities and timetable of the evaluation that was developed in step 12 should be consulted to ensure that the budget takes into account all the activities planned for the implementation of the evaluation.

If the evaluation is being co-financed by several organisations, it is useful to show the sources of funding to ensure that the funds are available when the evaluation is ready to be launched. If the evaluation is being co-sponsored without the transfer of funds—that is, the contributions are “in-kind”—then, these contributions can be described in the budget if they were not already listed in step 13 as part of roles and responsibilities. In any case, it is helpful to have all information in one place. For example, one organisation may provide transportation, another office space and secretarial services or a consultant. With the latter, it is important to make sure in accepting the offer of consultants that their skill-mix is appropriate for the specified needs of the evaluation.

The sponsoring organisation will use its own budgeting format and rules in preparing a budget. If the evaluation is to be outsourced, the bidding documentation would include a budget page where the total proposed budget for the evaluation should be indicated

with the budget items in the desired format. The bidders are usually asked to prepare and submit their best estimate of costs, both indirect (overhead and fringe benefits, if appropriate) and direct costs. A generic costing breakdown that might be useful is presented in **Table 10**.

**Table 10: Illustrative cost categories**

<b>Cost category</b>	<b>Budget</b>	<b>In-kind costs</b>
Personnel (daily rate)		
Fringe benefits		
Transportation		
Per diem		
Office supplies		
Telephone/Fax/Internet		
Rent and furnishings		
Repairs and maintenance		
Other		
Overhead or indirect costs		
Total		



# Conclusion

Prevention evaluation helps answer fundamental questions about trends in the HIV epidemic and the effectiveness of the response. Comparing evaluation results of multiple programmes and conducting evaluations across programmes helps stakeholders determine the overall impact of their efforts on populations.

Evaluations are also conducted to determine the merit and worth of a specific intervention or programme; they seek to determine what is working, what is not working, and why. They enable stakeholders to measure and document whether a programme has achieved its intended results. Based on evaluation results, decision makers can draw lessons and make programmatic recommendations for the future.

In prevention evaluation, evaluators face the difficult task of applying rigorous scientific standards within diverse geographic and population settings. Because evaluators make judgments as part of analyzing and interpreting data—and determining their implications for policy and programming—the evaluation community of practice also abides by ethical principles put forward by scientific institutions and professional associations. These scientific standards and ethical principles support the systematic development of an evaluation framework into what is called terms of reference (TOR).

TOR are an essential part of the planning process for a prevention evaluation. Working with a TOR template, multiple stakeholders can create a shared understanding of the specific purposes for the evaluation, the evaluation design and data collection needs, the resources available, the roles and responsibilities of different organisations, the timelines, and other fundamental aspects of the evaluation. Developing a TOR facilitates the development of the detailed evaluation study protocol that will then guide the implementation of the evaluation. The TOR is the foundation for robust design, effective implementation, rigorous data analysis and timely dissemination. Fidelity in following the TOR and the evaluation study protocol is essential in fulfilling the purpose of the evaluation. If changes need to be made (consistent with the overall objectives of the evaluation) due to unforeseen circumstances, they need to be documented and assessed for their impact on the evaluation process and the evaluation results.

Evaluations are conducted to learn, to better target programmes and reach more of the people for whom they are intended, and to make programmes more effective and efficient. They contribute to better management, governance, and learning. By using TOR to underpin evaluation protocols, evaluations can be more rigorous, the process of conducting evaluations and their relationship to policy and programme improvement can be better understood, and comparisons of programmes can produce more meaningful results. More systematic, rigorous, and credible evaluations are greatly needed in the fight against HIV, and a shared understanding of their purposes and processes will help to expand their usage. A sound evaluation TOR are the first critical step towards this result.

# Appendix 1: Illustrative Evaluation TOR Using the Template

This section presents a fully developed TOR to show how the 15 steps of the TOR template discussed in the guidance work together. It is fictional but realistic example. In this particular case, the director of the national AIDS programme has asked the M&E unit to prepare TOR that will be used for the preparation of a request for proposals that will be sent out to pre-qualified organisations that have demonstrated capacity to design and carry out evaluations. The type of evaluation sought is a combination of process and outcome evaluation. It will be planned and carried out with the Ministry of Health (MOH).

This illustration shows how the 15 sections can be abbreviated, combined with other sections, or expanded to include additional items not specified in the standard TOR template. It shows how the TOR template is a guide, that it can be easily adapted to address local needs without losing its rigour. It is, in other words, a tool that is flexible and can be adapted to a variety of situations.

## Sample: Fully developed TOR

### 1. Evaluation Topic

The MOH believes its pilot behaviour change communication / HIV counseling and testing (BCC/HCT) prevention project is successful and would like to expand it. However, there are no data to corroborate this. The MOH and the M&E unit believe an evaluation is needed to determine whether this project is effective or not, and whether it is cost-effective enough to be expanded.

### 2. Background

The MOH has been fighting a concentrated HIV epidemic for the past decade. The epidemic has been concentrated in one border area in the north of the country. HIV prevalence among the general population is less than 1 percent. However, incidence figures from the MOH show that along the northern border the number of new cases of HIV went from 7.7 percent in 2005 and to 12.0 percent in 2007. Recent data show that most-at-risk populations are driving these increases. Men who have sex with men (MSM) have a HIV prevalence of 9 percent and female sex workers (FSW) 1 percent. Sexual transmission of HIV was the most common form of transmission in 2007 (94 percent); perinatal transmission accounted for 4 percent.

The 2006 National Health and Demographic Survey (DHS) did not report incidence or prevalence but it did show that 98 percent of women aged 15 to 49 years had heard of AIDS and 97 percent knew of at least one way to avoid HIV infection. As might be expected, knowledge was highest among the more educated and those living in urban areas.

In 2006, the MOH introduced a pilot project to address the rapidly increasing expansion of HIV along the northern border. The project targeted most-at-risk populations with outreach, education, provision of HIV counseling and testing (HCT), distribution of free condoms, and a limited amount of social marketing of condoms. Unfortunately, there was no M&E unit at the time and only a limited amount of data has been collected about the

performance of this pilot project. Most of these data are routine service statistics (such as number of visits, contacts, condoms distributed, etc.). No HIV prevalence survey is planned until 2010, but there are some baseline data.

### 3. Evaluation purpose

The objective is to determine whether the expected results in changes in utilization of services and risk behaviour have been achieved among the targeted populations.

Interventions include:

- Behavioral change communication (BCC) for most-at-risk populations
- HIV counseling and testing (HTC) for most-at-risk populations
- Mass media aimed at most-at-risk populations
- Condom distribution at truck stops

### 4. Users

The principal user of this evaluation will be the MOH, and in particular, the AIDS Directorate and the M&E unit. There are quite a few stakeholders including implementing partners that are interested in the results (e.g., key donor agencies that are willing to support expansion of the project if it is achieving good results). The evaluation findings will be used to decide whether the programme should be expanded.

### 5. Key evaluation questions

The purpose is to *determine whether the expected results in changes in utilization of services and risk behaviour have been achieved among the targeted most-at-risk populations*. In addition, there are a number of other questions that the evaluation team should address.

#### Objective 1: Has the pilot programme achieved its expected results?

- What are the strengths and weaknesses of the project?
- What factors (internal and external) contributed to achievements?
- What lessons have been learned that should be applied to the pilot now and to the expansion (if that occurs) later?
- What kind of support is there for this pilot among most-at-risk populations, the general population, nongovernmental organizations, and current staff?
- What recommendations would the team make for the future?

#### Objective 2: How much has the pilot project cost?

- What are the costs by line item (personnel, transportation, etc.)?
- Have there been changes over time—if so, what, and what were the causes?
- What have been the financial advantages of the approach used?
- What could be done in the future to reduce costs without affecting results?
- Develop a financial analysis of the costs of the pilot project and an estimate of the additional costs of an expansion.

### Objective 3: Assess the costs and benefits of each of the prevention interventions

- What are the relative strengths and weaknesses of each intervention?
- What priority should be given to each of these interventions in the future, and why?
- Might other interventions be more effective and less costly?

## 6. Target groups

The evaluation will concentrate on most-at-risk populations in the designated intervention area along the northern border. Given the relative contribution of each of these groups to the spread of HIV, the first priority should be men who have sex with men (MSM), followed by female sex workers (FSW) and truck drivers at selected truck stops in the pilot area. The evaluation needs to make sure to sample these three populations adequately. Both qualitative as well as quantitative data will need to be collected and analyzed for these groups.

Secondary target groups are wives of truck drivers, other clients of FSWs, IDUs, pregnant women, out-of-school youth, and any other significant risk group identified.

## 7. Prevention interventions

As mentioned previously, there are four principal interventions:

- **Behavioral Change Communication (BCC)** for most-at-risk populations: This is an inter-personal peer education intervention. Twenty peers make periodic visits to areas where the target populations are known to congregate. They spend between one-half and one hour at each site to provide information and suggestions for ways to avoid contracting HIV.
- **HIV Counseling and Testing (HCT)** for most-at-risk populations: The MOH has scheduled 10 to 15 sessions each month for HIV counseling and rapid testing, usually based on recommendations from, and often with, peer educators participating in the sessions.
- **Mass media aimed at most-at-risk populations:** a regional nongovernmental organisation (Stop AIDS) sponsors various media buys (radio, local TV, community events, etc.) targeted at most-at-risk populations and the general public in the target areas.
- **Condom distribution at truck stops:** The Stop AIDS organisation also provides free condoms at truck rest stops and directly to female sex workers in those areas.

Each of these interventions should be mapped-out in a results chain to make explicit the assumptions and expected causal changes. Each intervention can be examined in two ways: first, to determine how the intervention has been carried out (e.g., did it do what it planned to do?); and second, to look at the results achieved against the baselines (e.g., did the project achieve the outcomes expected? How successful has the project been in changing the behaviour of each of the targeted populations?).

It will be important to describe the implementation process for each intervention so that an appropriate design can be selected for the evaluation. Cost data will be also be impor-

tant as it is an element needed in the decision making process regarding potential programme expansion.

## 8. Prevention indicators

The team should list and examine each of the key indicators of the project and their baseline values. The following key indicators will be measured:

### a. Outcome: Knowledge

- Percentage of most-at-risk populations who both correctly identify ways of preventing the sexual transmission of HIV and who reject major misconceptions about HIV transmission.

### b. Outcome: Behaviour

- Percentage of female sex workers reporting use of a condom with their most recent client.
- Percentage of men reporting use of a condom the last time they had anal sex with a male.

### c. Output: Activity

- Number of people in each most-at-risk population contacted
- Number of people in each most-at-risk population tested and counseled for HIV
- Number and percent of pregnant women tested for HIV
- Number of condoms distributed
- Total and average cost of the interventions

## 9–11. Methodology: Evaluation design, key data sources, and analysis procedures

These sections may be summarized as follows: A combination of a quasi-experimental design and a non-experimental design is planned; the key data sources would include:

- A **desk review** will be conducted of all relevant documents, including annual/quarterly work plans, progress reports, research reports and M&E reports.
- A **structured survey** of the targeted most-at-risk populations. Given the lack of hard outcome data for the last two years, a retrospective, representative survey of approximately 1,500 respondents will be used to collect needed data on key indicators.
- **Health service statistics** from all 20 district health centers and six hospitals will provide basic input and output data, including costs, for most project activities.
- A **qualitative survey** of approximately 200 individuals of most-at-risk populations, health staff, and NGOs will address the qualitative aspects of the evaluation.

Note: The elaboration of the above sections will be done in the evaluation study protocol. The remaining sections of the TOR need to be as specific as possible, but the content may vary. If the evaluation is to be outsourced, the TOR include what is expected from the contractors/consultants. If the evaluation is to be done by an internal team, these sections need to be equally specific but taking into consideration that those individuals asked to lead and/or participate in the evaluation will need to relinquish other obligations to conduct the evaluation, and others need to be hired to take on those responsibilities.

## 12. Evaluation activities and schedule

The time needed to complete an evaluation varies greatly with number of evaluation questions, sample size, locations, and other factors. The example used here could take between three and six months. The TOR will include a chart or table that lists the key activities as well as their estimated start and end dates.

The TOR will include what the expectations are from those conducting the evaluation whether it is done internally or by consultants. If the evaluation is outsourced, bidding documentation and consultants' TOR can include more specificity, such as the information below.

The following key deliverables (with delivery dates) should be highlighted in the TOR: a work plan, an initial briefing, a mid-term debriefing, a final debriefing, a preliminary summary of key findings, conclusions and recommendations; a first draft of the report; and a final draft of the report. Any changes in these deliverables must be approved by the M&E unit. Validation and dissemination workshops should also be considered.

The proposed final report should be no more than 40 pages, excluding annexes. The executive summary should be no longer than four pages (a rule of thumb is about 10 percent of the text length). The outline of the report should include a title page, abbreviations, a table of contents, an executive summary, introduction/background, evaluation design, data collection methods, findings, conclusions, lessons learned, and recommendations. In fact, the outline of the TOR can be used as a guide for the preparation of the report. Annexes should include the TOR, lists of the people contacted and documents reviewed, and any other annexes deemed important by the evaluation team.

## 13. Evaluation team members and level of effort (LOE)

The TOR need to include the information related to the skill-mix and resources needed to carry out the evaluation. This information can be used for contracting or for making administrative decisions if the evaluation is done in-house. For instance, the TOR may include the following:

- The evaluation team will consist of a team leader, a BCC specialist, a HIV specialist, a statistician, an M&E specialist, a budget specialist, and data collecting and data entry personnel. Not all specialists may be needed full-time. It is expected that all specialists will be recruited locally. No international consultants are anticipated.
- The structured survey will require approximately 10 interviewers for one to two weeks, depending on the sites selected. A survey specialist may be needed for one to two weeks if the HIV specialist does not have survey experience.
- The total level of effort required for the evaluation is estimated to be 450 person-days (see the budget for details).
- Technical personnel are expected to have a track record in their areas, an advanced degree in a health-related field or advanced technical skills, and experience in the AIDS arena. The senior specialist and team leader needs to demonstrate excellent skills in written and spoken language (English and French). The team leader also needs to show ability to coordinate and work closely with other team members.

Note: Remember that even if the evaluation is not done in-house, there still needs to be a manager responsible for managing the evaluation and the evaluation team advising (i.e., the different specialists). Their roles and responsibilities need to be spelled out. The same applies to co-sponsors of the evaluation.

**Role and qualifications of the team leader:** This person will have overall responsibility for the design, execution, and reporting of the evaluation. He/she will make assignments, review drafts, prepare briefings, and finalize the report. He/she should have at least seven years experience in AIDS and at least three assignments as a team leader. He/she will report directly to the M&E unit director.

**Role and qualifications of the BCC specialist:** This person will have primary responsibility for the sections of the evaluation that deal with BCC. He/she should have at least five years experience in AIDS and at least three assignments as a BCC specialist.

**Role and qualifications of the AIDS specialist:** This person will have primary responsibility for the data collection, including the surveys. He/she should have at least five years experience in AIDS and at least three assignments as an AIDS and/or survey specialist.

**Role and qualifications of the financial specialist:** This person will have primary responsibility for dealing with cost data and cost projections. He/she should have at least five years experience in health financing, especially in the area of AIDS.

**Role and qualifications of the M&E specialist:** This person will work with or oversee the individuals developing the data instruments and surveys. He/she should have at least five years experience in M&E in health, particularly in the area of AIDS.

**Role of the survey specialist:** This person will work with the other specialists to design, conduct and report on the structured surveys and qualitative data collection methods. He/she should have at least five years experience in survey work, especially in the area of AIDS.

## 14. Administrative & logistical support

The MOH will provide a contract to the successful bidder that will provide funds for all direct and indirect costs with the following exceptions: meeting space, office space, photocopying and printing, and office equipment – these items will be provided directly by the MOH.

## 15. Evaluation budget (cost-sensitive material — do not divulge)

The estimated cost for this evaluation is \$53,500. The breakdown by budget line item is shown in **Table 11**. Bidders are expected to prepare their own estimate of cost and include it in their submission. The actual amounts shown here are for illustrative purposes only.

**Table 11: Illustrative estimated cost of the evaluation**

Cost category	Units 1	Units 2	Unit cost	Total cost
Team leader (days)	100		\$50	\$5,000
BCC specialist (days)	100		\$40	\$5,000
AIDS specialist (days)	100		\$40	\$5,000
Budgeting specialist (days)	50		\$40	\$2,000
Survey specialist (days)	50		\$40	\$2,000
Interviewers (days * people)	10	10	\$10	\$1,000
Subtotal personnel (days)	450			\$20,000
Overhead @ 25%	25%		\$20,000	\$5,000
International travel	0		0	0
Local travel (trips)	500		\$5	\$2,500
Hotel (days, persons)	10	15	\$15	\$2,250
Subtotal				\$50,000
Indirect cost if appropriate (7%)				\$3,500
Total				\$53,500



# Appendix 2: Illustrative Summaries of Evaluation TOR

Above, we have illustrated a full application of the TOR in 15 steps. The following are examples of summary evaluation TOR that were compiled during the piloting of the TOR template in Nigeria. Two real evaluation studies were used.

## Illustration 1: Prevention outcome and impact evaluation

### Family Life HIV and AIDS Education (FLHE) curriculum implementation

1. Background <sup>8</sup>	The national response to HIV dates from 1986 when the first case of AIDS was identified, following which the National Expert Advisory Committee on AIDS (NEACA) was established in 1987. Since then, the HIV epidemic has steadily increased from 1.8% HIV prevalence in 1991, to 5.8% in 2001, and to 5% in 2003 and finally it retro-gressed to 4.4% in 2005. The population of people living with HIV is estimated to be 3.86 million, and AIDS-related deaths average 310,000 per year. Children living with HIV are of the order of 290,000 while children with one or both parents killed by AIDS-related conditions are believed to be as many as 1.8 million. The spread of HIV in the country has no boundary affecting both the private and public sector. The public sector in Nigeria undertakes essential functions and employs a significant proportion of those with technical skills and management expertise. Loss of human resources due to HIV is damaging to the public sector. The education sector, being unique in its role of training the young generation, will be seriously affected if the HIV epidemic is not brought under control. In 2002, the Federal Ministry of Education (FME)'s HIV/AIDS programme was initiated as part of the public sector programme with its focus on Family Life HIV and AIDS Education (FLHE). School-based baseline survey on HIV-related Knowledge, Attitudes, Practices, Skills (KAPS) & School Health in Nigeria, done in 2006, showed that a high proportion of the pupils engage in such risky behaviors as using salon clippers or needles (51%) and razor blades that had been used by others to cut their nails (33%). It is estimated that 20 new infections occur every day in Nigeria, 60% of which are in the 15–25 year old age group. The FLHE program for schools has the goal of reducing the HIV prevalence among in-school youth by 25% in five years.
2. Evaluation purpose	<ol style="list-style-type: none"> <li>1. To determine the effect of the intervention on the percentage of in-school youth aged 15–25 years who have not initiated sex.</li> <li>2. To determine the HIV incidence reduction in schools.</li> </ol>
3. Audience	Federal Ministry of Education, National AIDS Coordinating Authority, school management

<sup>8</sup> Note that another step (1. Evaluation Topic) was added after the workshop.

**Illustration 1: Prevention outcome and impact evaluation**

**Family Life HIV and AIDS Education (FLHE) curriculum implementation**

4. Key evaluation questions	What is the proportion of students aged 15–25 years who have not had sex? Has there been a decline in the incidence of HIV infection amongst students (15–25 years)?
5. Target group(s)	Secondary school students (15–25 years)
6. Prevention interventions	FLHE curriculum implementation — with focus on human development, personal skills, sexual health, relationships, sexual behaviour, society, and culture
7. Prevention indicators	<ol style="list-style-type: none"> <li>1. Percentage of schools with teachers who have been trained in life skills-based AIDS education and who taught it during the past academic year (output)</li> <li>2. Percentage of students aged 15–25 years who both correctly identify ways of preventing the sexual transmission of HIV and who reject major misconceptions about HIV transmission (knowledge outcome)</li> <li>3. Proportion of students aged 15–25 years who have not had sex (behaviour outcome)</li> <li>4. Percentage decline in the incidence of HIV infection among students aged 15–25 years (impact)</li> </ol>
8. Evaluation design	Non-experimental: pre- and post-test intervention. Assumption using ANC prevalence of 3.6% for age group 15–19, with population estimate of 10.3% of Nigerians (15–19). 15 million x 3.6% = 540,138 youth (15–19). Then 35% (students) of 540,138 = 189,048 secondary school student are HIV positive
9. Key data sources and procedures	<ol style="list-style-type: none"> <li>1. School-based survey on HIV-related Knowledge, Attitudes, Practices, Skills (KAPS) &amp; School Health (primary data source) using quantitative and qualitative procedures</li> <li>2. ANC sentinel survey technical report (2005 &amp; 2010)</li> <li>3. National population projection</li> <li>4. State Programmatic Quarterly Report (FLHE)</li> <li>5. Report of periodic supervisory visits</li> <li>6. NARHS (2005 &amp; 2007)</li> <li>7. Education Demographic Health Survey</li> </ol>
10. Key data analysis procedures	National & State level analysis—broken down by public and private schools; rural and urban, using SPSS. For cost-effectiveness the analysis model will calculate the costs as well as the effectiveness of the interventions.

### Illustration 1: Prevention outcome and impact evaluation

#### Family Life HIV and AIDS Education (FLHE) curriculum implementation

11. Evaluation activities and schedule	(1) Develop TOR (2) Identify various data sources (3) Develop an evaluation brief (4) Call for proposals to conduct survey (5) Development of evaluation tools (6) Selection of evaluators (contractor) (7) National and state level training (8) Field work for data collection and retrieval (9) Data entry and analysis (10) Report writing and production (11) Dissemination  *Estimated duration of evaluation activities = 6 months (Activity schedule to be attached.)
12. Team members and number of days	Epidemiologist, statistician, educationist (with HIV experience), Federal Ministry of Education, SOME, UNICEF & UNESCO  (Estimated number of days to be attached.)
13. Administrative & logistical support	(1) Office consumables (2) Communications (3) Travel cost (4) Accommodation & travel allowances (5) Meeting costs
14. Budget	*Estimated at N301,863,370 (based on historical figures)

### Illustration 2: Prevention outcome evaluation

#### National Behaviour Change Communication (BCC) Strategy for most-at-risk populations

1. Background <sup>9</sup>	Nigeria, the most populous African nation, has a population of 140 million people. With an adult HIV prevalence of 4.4 percent (2005, FMOH), the country has the world's third largest burden of HIV and AIDS after South Africa and India. The population of people living with HIV and AIDS is estimated to be 3.86 million, and AIDS-related deaths average 310,000 per year. In 2003, the National Agency for the Control of AIDS (NACA) developed a five-year national HIV and AIDS Behavior Change Communication Strategy 2004–2008 (BCCS) with technical support from the Johns Hopkins University Center for Communication Programs (JHU/CCP). The aim of the strategy was to coordinate and guide HIV and AIDS communication initiatives and messages among the various HIV and AIDS implementing partners in Nigeria. As the five-year period nears its completion, NACA is reviewing the appropriateness and effectiveness of the strategy in contributing to behaviour change with the aim of revising it in accordance with the latest research and lessons learned during the past four years and alignment with the National HIV and AIDS Prevention Plan 2007–2009.
----------------------------	---

<sup>9</sup> Note that another step (1. Evaluation Topic) was added after the workshop.

### Illustration 2: Prevention outcome evaluation

#### National Behaviour Change Communication (BCC) Strategy for most-at-risk populations

2. Evaluation purpose	To assess the effectiveness of national BCC strategies in reducing high-risk behaviour among most-at-risk populations in Nigeria over the period of 2004 to 2008. A sample will be drawn from six states (Abuja, Lagos, Cross River, Benue, Niger, and Kano).
3. Audience	National Agency for the Control of AIDS (NACA), donors, implementing partners.
4. Key evaluation questions	<ol style="list-style-type: none"> <li>1. Has risky behaviour been reduced among most-at-risk populations over the last four years?</li> <li>2. Which type of risky behaviour has changed?</li> <li>3. What are the characteristics or typology of the BCC strategy used?</li> <li>4. Is the reduction due to the strategy employed?</li> <li>5. Which one of the strategies contributed the most to the reduction of risky behaviour among most-at-risk populations?</li> <li>6. Which of the strategies is most cost-effective?</li> </ol>
5. Target group(s)	Female sex workers, uniformed services workers, transport workers, people living with HIV
6. Prevention interventions	Interpersonal communication, peer education; condom distribution, mass media; community awareness.
7. Prevention indicators	<ol style="list-style-type: none"> <li>1. Percent increase in condom use among most-at-risk populations</li> <li>2. Percent decrease in number of non-spousal multiple sex partners</li> <li>3. Percent increase of sex workers who in the past 12 months used a condom consistently during sexual intercourse with a client</li> </ol>
8. Evaluation design	Non-experimental time series
9. Key data sources and procedures	NARHS, BSS, IBBSS, ANC, cross-sectional surveys, programme monitoring data, new primary data (quantitative/qualitative studies)
10. Key data analysis procedures	Performance assessments, cost-effectiveness analysis, and response analysis
11. Evaluation activities and schedule	Three months, February to April (including preparation, field work, analysis and report)

**Illustration 2: Prevention outcome evaluation**

**National Behaviour Change Communication (BCC) Strategy for most-at-risk populations**

12. Team members and number of days	Sociologist, economist, epidemiologist, public health specialist
13. Administrative & logistical support	Secretariat support, hotel accommodation, transport, and incidental expenses will be provided by NACA, and the contractor will be responsible for all other support.
14. Budget	Estimated budget at \$350,000.

# Appendix 3: Supporting Information

## A. Evaluation Approach

An important part of evaluation planning is defining the conceptual approach that will guide the evaluation such as the use of a results chain, as explained in the guidance. In this section, we present an approach that incorporates the elements of a good evaluation and the results chain that enables us to identify what will be evaluated. Inherent in the application of this approach are the guiding principles that serve as the cornerstone of good and ethical evaluation practice.<sup>10</sup> The terms of reference (TOR) for prevention evaluation are, thus, placed within the context of accepted international standards for scientific study. (see **Section C**).

“The community of scientists is bound by a set of values, traditions, and standards that embody honesty, integrity, objectivity, and collegiality. These values are reflected in the particular principles and practices characteristic of specific scientific disciplines.”

[Source: Panel on Scientific Responsibility and the Conduct of Research. Committee on Science, Engineering and Public Policy. National Academy of Sciences. *Responsible Science: Ensuring the Integrity of the Research Process, Volume 1*. Washington, DC: National Academy Press, 1992,p.1.]

### Elements of Good Evaluation

Reliable and useful evaluations must meet basic quality standards and criteria.<sup>11</sup> TOR should incorporate all of these elements so that evaluations are credible. Credible evaluations are:

- **Impartial** — they are objective and free of bias.
- **Systematic and technically adequate** — they use sound methods of inquiry and follow a logical procedural model. To the extent possible and applicable, evaluations should be planned and undertaken with the rigour associated with more formal studies.
- **Valuable and useful** — they add value to management actions such as strategy design, selectivity, resource allocation, and programme implementation by providing credible strategic information in a cost-effective manner.
- **User-owned** — stakeholders are included in the design, planning and implementation of the evaluation. If evaluation findings are to be used for policy and programme improvements, then the users must have a stake in the evaluation and its findings and apply them so as to achieve better results, to lower costs, and to improve programmes.

10 The *Guiding Principles for Evaluators* were developed in 1994 as guidelines for sound, ethical practice and have been broadly vetted with the American Evaluation Association membership and reviewed and revised at regular intervals, including most recently in 2003, in order to ensure that they remain current with the field. The African Evaluation Association also has issued ethical principles for evaluators. See Part III Section A.

11 Adapted from R. Rodriguez-Garcia and E.M. White Self-Assessment in Managing for Results, World Bank Working Paper # 41, 2005, and Kusek and Rist, Ten Steps to a Results-based M&R System, The World Bank, 2004.

### **Feedback and dissemination**

They enable findings to be used for policy and programme improvements. In addition, evaluators pay special attention to the analytical phase where judgments are made about the findings and their implications. But how does one decide whether a response is well-reasoned? Evaluations are a tool for governance and transparency as they provide feedback on the outcomes and consequences of government policies and actions. Nongovernmental and civil society organisations include evaluations in their learning activities to gain information on programme results and the achievement of the organisation's mission. They report to stakeholders on achievements. Thus, how evaluation findings are judged by evaluators have significant implications for accountability and for the population organisations serve. Thus, evaluations need to be grounded in solid ethical principles of evaluation practice.

## B. Evaluation Designs

Design Elements	Type of questions that can be answered with this method	Complexity of design	Cost/time and resources	Ethical considerations	Potential for comparison group	Other considerations
<b>Process Evaluation</b>						
	<ul style="list-style-type: none"> <li>The main question: How can we improve or strengthen the programme under study?</li> <li>Is an intervention type feasible (or is an intervention modified for a different audience/ setting feasible)?</li> </ul>	<p>Simple design can use qualitative and/or quantitative methodologies</p> <p>Data sources include programme monitoring data, exit interviews, participant interviews, direct observation, cost data, provider (or facilitator) interviews; focus group discussions and key informant interviews.</p>	<p>Rapid assessments possible (e.g., one to three months)</p> <p>Inexpensive</p>	<p>For direct observations or interviews with participants or facilitators, evaluators need to obtain consent (verbal or written).</p> <p>Although cost data and programme monitoring data do not require consent, individual implementers may not always willingly release the information. Credibility on the part of the national authority is therefore key to obtaining these data.</p>	<p>Not applicable</p> <p>May be able to make comparisons on challenges and lessons learned across different sites</p>	<p>Process evaluations are useful for determining whether programmes are being implemented as planned and whether it is worth undertaking an outcome evaluation of the programme</p> <p>In many cases, it is crucial to obtain process information as well to determine if the programme is being implemented as planned.</p>
	<p>Is it possible to replicate an intervention?</p> <p>What were some of the challenges that were encountered during implementation?</p> <p>How useful were the guidelines or materials that were developed for the intervention?</p> <p>How much does an intervention cost?</p>					



Design Elements	Type of questions that can be answered with this method	Complexity of design	Cost/time and resources	Ethical considerations	Potential for comparison group	Other considerations
<b>Outcome Evaluation</b>						
<b>Non-experimental (pre- and post-test design)</b>	<p>Does the intervention have an effect on behaviors? Does the programme have an effect on norms?</p> <p>Does the programme have an effect on knowledge? Does the programme have an effect on attitudes or stigma?</p> <p>What is the coverage of an intervention (or of multiple interventions)?</p> <p>Does the intervention lead to program utilization?</p>	<p>No comparison or control group necessary — examine participants (or target community) at beginning and end of the programme and compare changes</p> <p><b>Note:</b> Multivariate analysis is required in order to control for the effect of multiple causal factors. Bias may arise if you do not have a representative sample.</p>	<p>Cost: less expensive, but this may vary from country to country</p> <p><b>Note:</b> Time series requires baseline data. Pre- and post-test requires follow-up period prior to the post-test.</p>	<p>Whenever interviewing individuals, there is a need to determine whether verbal or written consent is required. If blood or other specimens are taken, additional consent procedures are also required.</p> <p><b>Note:</b> Since this may involve the use of survey methodology, an appropriate national authority should review the protocol prior to conducting the evaluation</p>	<p>Not required</p> <p>Possibility for attribution to the programme or intervention is very weak.</p>	<p>Non-experimental designs are suggestive of positive changes and cannot eliminate the possibility that the observed change would not have happened anyway (i.e., even without the programme or intervention present).</p>

Design Elements	Type of questions that can be answered with this method	Complexity of design	Cost/time and resources	Ethical considerations	Potential for comparison group	Other considerations
<b>Quasi-experimental (with comparison group or time series design)</b>	Does the intervention have an effect on behaviours? Does the programme have an effect on norms? Does the programme have an effect on knowledge? Does the programme have an effect on attitudes or stigma? What is the coverage of an intervention (or of multiple interventions)? Does the intervention lead to program utilization?	Requires comparison group (for non-time series designs). Comparison group can be matched by certain characteristics (especially those that may independently affect outcomes). <b>Time series</b> design requires multiple observations from the same individuals or communities over time. May require multivariate analysis.	<b>Cost:</b> More expensive than non-experimental because includes individuals or communities that were not part of the intervention. <b>Time:</b> Requires baseline data and a follow-up period prior to the post-test and thus a longer period of data collection (before and after intervention).	Whenever interviewing individuals, there is a need to determine whether verbal or written consent is required. If blood or other specimens are taken, additional consent procedures are also required. <b>Note:</b> Since this may involve the use of survey methodology, an appropriate national authority should review the protocol prior to conducting the evaluation.	Requires a comparison group — not randomly assigned — could use a generic comparison group (e.g., from a national survey or another source). <b>Time series</b> does not require comparison group.	The best option for an outcome evaluation with the objective of determining if the programme led to change in the outcome is a quasi-experimental design but often complex to implement depending on the programme to be evaluated.

Design Elements	Type of questions that can be answered with this method	Complexity of design	Cost/time and resources	Ethical considerations	Potential for comparison group	Other considerations
<b>Impact Evaluation</b>						
Quasi-experimental	Did the programme(s) lead to a change in HIV incidence? Did the programme lead to a change in HIV prevalence? Did the programme lead to a change in HIV morbidity? Did the programme lead to a change in HIV mortality?	Requires comparison group (for non-time series designs). Comparison group can be matched by certain characteristics. <b>Time series</b> design requires multiple observations from the same individuals or communities over time. Requires multivariate analysis.	<b>Cost:</b> expensive because includes population-level data from individuals or communities that were not part of the intervention (surveillance data could be used). <b>Time:</b> Pre- and post-test require baseline data and a follow-up period prior to the post-test, and thus a longer period of data collection (before and after) is needed (surveillance data could be used).	Whenever interviewing individuals, there is a need to determine whether verbal or written consent is required. If blood or other specimens are taken, additional consent procedures are also required. <b>Note:</b> Since this may involve the use of survey methodology, an appropriate national authority should review the protocol prior to conducting the evaluation.	Requires a comparison group — not randomly assigned — could use a generic comparison group (e.g., from a national survey or another source).	The best option for an impact evaluation, if the objective is to show the programme had an impact, is often a quasi-experimental design but often complex to implement depending on the programme to be evaluated. <ul style="list-style-type: none"><li>Impact should be clearly defined and distinguished from proxy variables. The gold standard for the measurement of impact for an HIV prevention activity is HIV incidence (HIV prevalence among those ages 15–19 can be used as a proxy). Mortality, on the other hand, is usually the gold standard for overall HIV control activities.</li><li>It is usually difficult (and mostly unnecessary) to attribute impact to any single programme; often observed changes may be the combined effect of several programmes operating in an area.</li></ul>

Design Elements	Type of questions that can be answered with this method	Complexity of design	Cost/time and resources	Ethical considerations	Potential for comparison group	Other considerations
<b>Impact Evaluation</b>						
Experimental	<p>Did the programme(s) lead to a change in HIV incidence?</p> <p>Did the programme lead to a change in HIV prevalence?</p> <p>Did the programme lead to a change in HIV morbidity?</p> <p>Did the programme lead to a change in HIV mortality?</p>	Requires a randomly assigned control or comparison group (individuals or clusters/communities). May require multivariate analysis	<p>Cost: randomization of individuals or clusters/communities may make the evaluation more costly, depending on availability of control/comparison sites (individuals).</p>	<p>Whenever interviewing individuals, there is a need to determine whether verbal or written consent is required. If blood or other specimens are taken, additional consent procedures are also required.</p> <p>Note: Since this may involve the use of survey methodology, an appropriate national authority should review the protocol prior to conducting the evaluation</p>	Requires a control group that is randomly assigned at the individual or cluster/community level.	Experimental designs are often too costly and difficult to implement as randomization at the individual or cluster/community levels is often not considered appropriate or acceptable. If an experimental design is planned, adequate planning and resource mobilization needs to be undertaken prior to conducting the evaluation.

## C. International Guiding Principles for Evaluators

Guiding Principle	Related Issues and Questions
<p><b>Systematic Inquiry</b> How are methods of participant selection important to the credibility of an evaluation?  What about methods of data collection?</p>	<ul style="list-style-type: none"> <li>• How were neighbourhood residents as well as programme participants included in evaluation activities?</li> <li>• What procedures were established or used to ensure systematic inclusion of those stakeholders' perspectives?</li> <li>• How were focus group participants recruited (potential bias)? (There are related issues in sampling, field testing instruments, and survey design.)</li> <li>• In what ways does the evaluation address potential weaknesses of convenience in gathering data as opposed to systematic methods, documented in a work plan?</li> <li>• How well do the evaluation design and its questions address the information needs it is intended to meet?</li> </ul>
<p><b>Competence</b> How can you decide what dimensions of competence are relevant for an evaluation?  What is cultural competence and how will you know its presence?</p>	<ul style="list-style-type: none"> <li>• What elements of competence are brought into play and what omissions (in competence) might be envisioned?</li> <li>• How does cultural competence weigh with respect to (a) evaluation competence (evaluation knowledge, skills, experience — what do we assume from academic credentials and some relevant experience); (b) program competence (knowledge of policies, mission, staffing, and so on); and (c) social or environmental competence (understanding and appreciation of relevant social, economic, and political realities)?</li> <li>• For cultural competence, are language skills necessary? Are they sufficient?</li> <li>• In what ways might competence be usefully or realistically judged from the eyes of the buyer of evaluation services?</li> </ul>

Guiding Principle	Related Issues and Questions
<p><b>Integrity/Honesty</b> How might the idea of integrity or honesty affect different stages of the evaluation?</p>	<ul style="list-style-type: none"> <li>• What possible conflicts might arise from the evaluator's previous experience?</li> <li>• How can these be resolved or handled?</li> <li>• What potential influences might be created by multiple uses to which the evaluation might be put (e.g., academic publication, etc.) and how could these be addressed?</li> <li>• What mechanisms or agreements were (or might be) established to deal with potential conflicts or issues that might arise during the course of the evaluation?</li> <li>• Why wasn't a request for proposals (RFP) issued and/or a written proposal used as the basis for explicitly establishing mutual expectations for the evaluation?</li> <li>• How does this principle affect the inclusion of participants and residents (or ways in which they are included and supported) in the evaluation activities?</li> </ul>
<p><b>Respect for People</b> What are major ways of showing respect for people in an evaluation context?</p>	<ul style="list-style-type: none"> <li>• How does this principle interact with the need for cultural competence?</li> <li>• What about the apparently minimal inclusion of programme participants in the evaluation activities, and how about the lack of inclusion of neighbourhood residents who are not programme participants?</li> <li>• Is there institutional review board (IRB) approval, and does it ensure sufficient and appropriate respect for people (clients, staff, etc.) in the evaluation?</li> <li>• What issues of respect might arise regarding compensation for participation?</li> </ul>
<p><b>Responsibilities for General and Public Welfare</b> Are there limits to an evaluator's responsibilities for general and public welfare?</p>	<ul style="list-style-type: none"> <li>• How does the planned dissemination meet these responsibilities, or what issues are raised in possibly not meeting them?</li> <li>How might the evaluation results negatively impact the neighbourhood or residents?</li> <li>How might the evaluator (or those managing the evaluation) consider and address such risks?</li> <li>What pressures or influences might be anticipated from the power/status relationships at work in the evaluation concerning, for example, participants in relation to staff; both in relation to funder/ "owner" of the evaluation; between "owner" and funder?</li> </ul>

[Source: American Evaluation Association. Guiding Principles for Evaluators. [www.eval.org](http://www.eval.org)]

### Guidelines for Evaluation in Africa: The African Evaluation Association

These guidelines for evaluation (available at [www.afrea.org](http://www.afrea.org)) are ethical principles and quality criteria. They are a set of rules and requirements that are applicable to all stakeholders and throughout the evaluation process. These principles show a shared system of values among evaluation stakeholders in Africa. The evaluation standards help enhance independence and impartiality in the conduct of evaluation. They ensure transparency and a participative methodology and create conditions of ownership of evaluation and its results. Also, they aim to standardize methods and upgrade the trustworthiness and usefulness of evaluation. This second edition (2006) of the evaluation guidelines aggregates the various works of stakeholders in the field of evaluation in Africa. It is also in tune with the major trends and good practices in evaluation worldwide. The new edition of guidelines in evaluation both takes into account universal standards and promotes requirements justified by the state of evaluation in Africa. It supervises the process and products and embraces all sectors and time frames of the evaluation project. These standards target all types of evaluation. They are maximal and incorporate a pluralist dimension. They help give credit to evaluation stakeholders, make reliable the evaluation processes, and better professionalize evaluation in Africa. They include 34 rules divided into four major principles:

- **Utility:** for produced information and expected and provided results.
- **Feasibility:** for realism, cautiousness, and efficiency.
- **Respect of ethics:** for respect of legal and ethical rules.
- **Precision and quality:** for a relevant methodology related to the goal and the subject matter of the evaluation.

## D. Costs and Cost-effectiveness of Prevention Interventions

The process of estimating the cost per intervention is based on internationally agreed standards for costing. The following table is taken from a World Bank study undertaken in Honduras in 2002.<sup>12</sup> It is rare in that it is a cost-effectiveness evaluation. It shows the cost of an infection avoided by each of 12 prevention interventions. For example, condom social marketing avoids an infection for only US\$1.70. HIV counseling and testing costs \$18.29 per infection avoided. Prevention of mother-to-child transmission (PMTCT) is much more expensive at over \$630 per infection avoided. Because of the prevalence rate of around 2 percent and the pattern of infections, large numbers of pregnant women need to be tested to find a single infection, and approximately three HIV-positive women and their infants need to be treated to prevent one infection.

Evaluations of costs and cost effectiveness are important for decisions about whether to scale up, scale down, continue, drop or replicate different interventions. By considering the cost per infection, managers can allocate their funds across those interventions that provide the greatest return on investment, see **Table 12**.

**Table 12: Illustrative unit costs of prevention interventions**

<b>Intervention</b>	<b>Average Unit Cost US\$</b>
1. Condom social marketing.	\$1.70
2. Condom distribution in most-at-risk populations	\$3.59
3. Information, Education, Communication (IEC) for ethnic group — Garifunas	\$8.68
4. IEC for pregnant women	\$9.15
5. IEC for adolescents	\$10.44
6. IEC targeted at most-at-risk populations (sex workers, men who have sex with men, prisoners)	\$11.10
7. Counseling and rapid HIV testing	\$18.29
8. Supporting the promotion and defense of human rights	\$20.13
9. Interventions in the workplace	\$20.88
10. Blood safety	\$24.10
11. Syndromic management of sexually transmitted diseases	\$48.97
12. Strengthening the vertical transmission prevention programme	\$636.19



## E. Glossary of M&E Terminology

- **Activity** refers to actions taken or work performed through which inputs such as funds, technical assistance, and other types of resources are mobilized to produce specific outputs.
- **Baseline** provides information (quantitative or qualitative) that provides a value for an indicator at the beginning of, or just prior to, the monitoring period.
- **Effectiveness** measures the merit and worth of an intervention and the extent to which the intervention objectives were achieved or are expected to be achieved in a sustainable manner and with positive institutional effects.
- **Evaluation** is a rigorous, scientifically-based collection and analysis of information about programme activities, characteristics, and outcomes that determine the merit or worth of a specific programme. Evaluation studies are used to improve programmes and inform decisions about future resource allocations.
- **Goal** is a broad statement of a desired, usually longer-term, outcome of a programme/ intervention. Goals express general programme/intervention intentions and help guide the development of a programme/intervention. Each goal has a set of related, specific objectives that, if met, will collectively permit the achievement of the stated goal.
- **Impact** the long-term, cumulative effect of programmes/interventions over time on what they ultimately aim to change, such as a change in HIV infection, AIDS-related morbidity and mortality.
- **Impact evaluation** looks at the rise and fall of disease incidence and prevalence as a function of AIDS programmes. The effects (impact) on entire populations can seldom be attributed to a single programme or even several programmes, therefore, evaluations of impact on populations usually entail a rigorous evaluation design that includes the combined effects of a number of programmes.
- **Indicators** are quantitative or qualitative variables that provide valid and reliable ways to measure achievement, assess performance, or reflect changes connected to an intervention.
- **Inputs** include financial, human and material resources.
- **Intervention:** a specific activity (or set of activities) intended to bring about change in some aspect of the status of the target population (e.g., HIV risk reduction, improving the quality of services) using a common strategy. An intervention has distinct process and outcome objectives and a protocol outlining the elements and benchmarks of the intervention. Interventions are implemented by a project or providers and can be focused at various levels such as the individual, small or large group, community or societal levels.
- **Managing for Results** refers to a comprehensive and integrated management system that focuses on achieving national objectives for the population while assuring accountability for public funds.\*
- **Monitoring** is the routine tracking of key elements of a programme or project, its outputs and its intended outcomes.
- **Outcomes** are short-term and medium-term effect of an intervention's outputs, such as change in knowledge, attitudes, beliefs, behaviours.

- **Outcome evaluation** is a type of evaluation that is concerned with determining if, and by how much, programme activities or services achieved their intended outcomes. Whereas outcome monitoring is helpful and necessary in knowing whether outcomes were attained, outcome evaluation attempts to attribute observed change to the intervention tested, describe the extent or scope of programme outcomes, and indicate what might happen in the absence of the program. It is methodologically rigorous and requires a comparative element in design, such as a control or comparison group.
- **Outcome monitoring** is the tracking of variables that have been adopted as valid and reliable measures (i.e., indicators) of the desired programme/intervention outcomes. Outcome monitoring does not infer causality; changes in outcomes may be attributable to multiple factors, not just a specified programme/intervention.
- **Outputs** are the results of program/intervention activities; the direct products or deliverables of programme/intervention activities, such as the number of HIV counseling sessions completed, the number of people served, the number of condoms distributed.
- **Process evaluation** is a type of evaluation that focuses on programme implementation and uses largely qualitative methods to describe programme activities and perceptions, especially during the developmental stages and early implementation of a programme. It may also include some quantitative approaches, such as surveys about client satisfaction and perceptions about needs and services. In addition, it might provide understanding about the cultural, sociopolitical, legal, and economic contexts that affect a programme.
- **Programme** is an overarching national or sub-national response to a disease. A programme generally includes a set of interventions marshaled to attain specific global, regional, country, or subnational objectives; involves multiple activities that may cut across sectors, themes and/or geographic areas.
- **Results-based evaluation** is an assessment of a planned, ongoing, or completed intervention to determine its relevance, efficiency, effectiveness, impact and sustainability. The intent is to incorporate the findings and lessons learned into the decision-making process.<sup>13</sup>
- **Surveillance** is the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health.
- **Target** is the objective a programme/intervention is working towards, expressed as a measurable value; the desired value for an indicator at a particular point in time.
- **Triangulation** refers to the analysis and use of data from multiple sources obtained by different methods. Findings can be corroborated and the weakness (or bias) of any one method or data source can be compensated for by the strengths of another, thereby increasing the validity and reliability of the results.

[Source: Glossary of M&E Terminology. UNAIDS Monitoring and Evaluation Reference Group June 2008; except for terms indicated by \* which source is: Organization for Economic Cooperation and Development, Development Assistance Committee, 2002.]

# References

- Adamchak S, Reynolds J, Henn J. *Assessment of Monitoring and Evaluation in Projects Managed by the Bureau for Global Health*. Poptech: Office of Population and Reproductive Health, 2004.
- African Evaluation Association (AfrEA). *Guidelines for Evaluation in Africa, 2006*. [www.afrea.org](http://www.afrea.org)
- American Evaluation Association. *Guiding Principles for Evaluations*. [www.aea.org](http://www.aea.org)
- Canadian International Development Agency. *How to Perform Evaluations. Model TOR. No2, 2000*.
- Kusek J, Rist R, *Ten Steps to a Results-based M&R System*. The World Bank, 2004.
- Quinn-Patton M. *Utilization-focused evaluation. 4th Edition*. Sage Publications, 2008.
- Reynolds J et al. *Evaluation Handbook for Family Planning Programs*, U.S. Department of Health, Education and Welfare, 1975.
- Rodriguez-Garcia R, White E. *Self-Assessment in Managing for Results*, World Bank Working Paper # 41, 2005.
- Rodriguez-Garcia R, Kusek J. *Planning and Managing for Results: A Handbook*. World Bank. Global AIDS Monitoring and Evaluation Team, World Bank Global HIV/AIDS Program, 2007.
- Rugg D, Peersman G, Carael M. (Eds.). *Global advances in HIV/AIDS monitoring and evaluation. New Directions for Evaluation 103, 2004*.
- UNAIDS. *Setting National Targets for Moving towards Universal Access: A working document, 2006*.
- UNAIDS. *Practical Guidelines for Intensifying HIV Prevention. Towards Universal Access, 2007*.
- UNAIDS. *A Framework for Monitoring and Evaluating HIV Prevention Programmes for Most-At-Risk Populations, 2007*.
- UNAIDS. *A Framework for Classifying HIV Prevention Interventions. A Report to the Joint United Nations Programme on HIV/AIDS by Michael Sweat, 2008*.
- UNAIDS. *Report on the Global AIDS Epidemic, 2008*.
- UNAIDS. *Glossary of M&E Terminology. UNAIDS Monitoring and Evaluation Reference Group, 2009*.

- United Nations. United Nations High level event on the Millennium Development Goals. Background Note by the Secretary-General, September 25, 2008.
- World Bank. *Optimizing the Allocation of Resources among HIV Prevention Interventions in Honduras*. World Bank. Washington, DC, July 2002.
- World Bank. *An Evaluation of Bank Support for Decentralization in Client Countries*, Fast Track Brief, 2008.
- World Bank GHAP. *Health Systems and HIV/AIDS: Impact on Service Delivery (working draft)*, 2008.
- World Health Organization. *Preventing HIV/AIDS in Young People: A Systematic Review of the Evidence from Developing Countries*. WHO Technical Report Series 938. WHO. Geneva, 2006.

## Electronic Sources Consulted

- African Evaluation Association: <http://www.afrea.org>
- American Evaluation Association: [www.eval.org/resources.asp](http://www.eval.org/resources.asp)
- Centers for Disease Control and Prevention (CDC): [www.cdc.gov](http://www.cdc.gov)
- Center for Disease Control and Prevention: Tiers of Evidence: A Framework for Classifying HIV Behavioral Interventions: [www.cdc.gov/hiv/topics/research/prs/tiers-of-evidence.htm](http://www.cdc.gov/hiv/topics/research/prs/tiers-of-evidence.htm)
- Global Fund to Fight AIDS, Tuberculosis and Malaria: [www.theglobalfund.org/](http://www.theglobalfund.org/)
- Measure Evaluation: [www.measureprogram.org](http://www.measureprogram.org) and <http://www.cpc.unc.edu/measure>
- UNAIDS: [www.unaids.org](http://www.unaids.org)
- United Nations Development Programs: [www.undp.org](http://www.undp.org) and [http://www.undp.org/gef/undp-gef monitoring evaluation](http://www.undp.org/gef/undp-gef_monitoring_evaluation)
- United States Government: <http://www.globalHIVevaluation.org>
- World Bank Global AIDS Program: [www.worldbank.org/aids](http://www.worldbank.org/aids) and [www.worldbank.org/asap](http://www.worldbank.org/asap)
- World Bank Independent Evaluation Group: [http://www.worldbank.org/oed/ecd/me\\_tools\\_and\\_approaches.html](http://www.worldbank.org/oed/ecd/me_tools_and_approaches.html)
- World Bank Institute: [www.worldbank.org/wbi](http://www.worldbank.org/wbi) and [www.worldbank.org/wbi/evaluation](http://www.worldbank.org/wbi/evaluation)



## List of ALL MERG Documents 2007-2009

- 1. A Framework for Monitoring and Evaluating HIV Prevention Programmes for Most-At-Risk Populations (2007):**  
Provides an overview of M&E methods and approaches for most at-risk populations; it covers the use of strategic information for programme planning, M&E. Its focus is on the M&E of targeted HIV prevention programmeM&E. Its focus is on the M&E of targeted HIV prevention programme
- 2. Additional Recommended Indicators. Addendum to UNGASS Monitoring the Declaration of Commitment on HIV/AIDS, Guidelines on Construction of Core Indicators (2008):** Presents the 40 core national indicators that provide minimum necessary information for national-level monitoring of the HIV epidemic and response, and to provide detailed specifications and guidance on the 15 indicators recommended in addition to the 25 UNGASS indicators
- 3. Organizing Framework for a Functional National HIV M&E System (2008):** This framework describes 12 main M&E system components and defines a performance goal and results for each component. The framework helps countries to define an agreed set of national performance objectives and measures for the HIV M&E system and to guide strategies for building capacity, where needed, to reach these objectives.
- 4. Glossary of M&E Terminology (2008):** contains an alphabetical listing of M&E terms and their definitions often with more in-depth explanations than would customarily be provided by dictionary definitions. The Glossary will facilitate and improve dialogue and understanding among all those who are involved in M&E of development activities. It should also serve as a valuable reference guide in M&E training. The selection of terms and their definitions in the attached glossary have been carefully discussed and endorsed by the Global UNAIDS Monitoring and Evaluation Reference Group (MERG)
- 5. Indicator Standards and Assessment Tool (2009):** consists of a set of agreed indicator standards that are relevant at the national level for program managers and service providers, who need to select, revise and use indicators to monitor, manage and implement their country's response to the epidemic effectively monitor. This will ensure that indicators provide decision-makers and key stakeholders with useful, feasible and relevant information. An additional aim is to reduce the burden of global reporting on countries by harmonising global level indicators across multilateral and bilateral organisations
- 6. Planning Tool for Developing a Digital Library of M&E Resources (2009):** A Planning Tool to help assure that users of a digital library can successfully locate resources and can make informed decisions regarding the quality of the materials. The Planning Tool has two purposes: 1) To provide guidance to current owners and future developers of a digital library on the range of issues to be addressed in usability and user-friendliness of the library and 2) To provide a list of questions to help organizations brainstorm if they can and should invest their resources in developing a digital library
- 7. Guidance HIV Monitoring and Evaluation Capacity-building (2009):** provides practical advice for national AIDS programmes that are planning and implementing capacity building activities as part of their effort to develop a unified and effective national HIV monitoring and evaluation (M&E) system. The Guidance is relevant to the wide range of individuals and organisations involved in the national HIV M&E system; it is particularly relevant for the health sector, given its central role in M&E of HIV.
- 8. 12 Components Monitoring and Evaluation System Assessment – Guidelines to support preparation, implementation and follow-up activities (2009):** These Guidelines provide information on the preparation for and implementation of an assessment of the national HIV monitoring and evaluation (M&E) system. It also includes key steps to take after an assessment to facilitate implementation of M&E system strengthening activities. The Guidelines are built around the 12 main components of the HIV M&E system, which define the Organizing Framework for a Functional National HIV Monitoring and Evaluation System (UNAIDS, 2008). Consequently, the Guidelines also focus on using the 12 Components Monitoring and Evaluation System Strengthening Tool (Geneva: UNAIDS, 2009a) to ensure a comprehensive and successful assessment.
- 9. 12 Components Monitoring and Evaluation System Strengthening Tool (2009):** Is a tool is for assessing how well each of the 12 components of a national HIV M&E system is functioning. The tool facilitates the identification of strengths and weaknesses in the national HIV M&E system and the prioritization of system strengthening activities.
- 10. Guidelines for Developing Terms of Reference for Prevention Evaluation (2009):** The Guidelines aim to foster a systematic approach to the evaluation of prevention programs by focusing on an often overlooked yet critical step in evaluation planning: the preparation of terms of reference (TOR). It can be used to facilitate the planning of evaluations for HIV prevention, discussions on the design of these evaluations, and the drafting of TOR to guide such assessments. It is intended for use by anyone who prepares or reviews TOR for evaluations of HIV and AIDS prevention programs and projects.



**UNAIDS**  
JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS

UNHCR  
UNICEF  
WFP  
UNDP  
UNFPA  
UNODC  
ILO  
UNESCO  
WHO  
WORLD BANK

UNAIDS  
20 AVENUE APPIA  
CH-1211 GENEVA 27  
SWITZERLAND

T (+41) 22 791 36 66  
F (+41) 22 791 41 87

[www.unaids.org](http://www.unaids.org)

Uniting the world against **AIDS**