GLOBAL AIDS PROGRAM

Monitoring & Evaluation Capacity Building for Program Improvement

Field Guide

Version 1

December 2003

IMPORTANT DISCLAIMER – READ FIRST

This Guide focuses on issues related to monitoring and evaluation for program improvement that can be applied in any setting, and to any program. However, *Chapter 13. GAP Planning, Budget, and Reporting System* and associated Appendices K, L and M sets out GAP-specific planning and reporting requirements for the reporting round completed in 2003. Note that these requirements may be updated for future reporting rounds, hence GAP offices should refer to guidance sent out from the GAP Office of the Director each year to meet their reporting requirements. The issues and templates in this guide are included for illustrative purposes only. Note also that this Field Guide was developed before the implementation of President Bush's Emergency Plan for AIDS Relief, hence it does not include any information on this new Initiative.

CITING AND ADAPTING INFORMATION FROM THIS GUIDE

You are encouraged to freely use materials contained in this guide. When directly quoting or copying materials from this guide, please use the following citation:

Global AIDS Program. (2003). *Monitoring and Evaluation Capacity Building for Program Improvement Field Guide, Version 1*. U.S. Centers for Disease Control and Prevention, Atlanta, GA.

When adapting portions of the guide, but not using quotes or material exactly as they are presented, please use the following disclaimer:

Adapted from: Global AIDS Program. (2003). *Monitoring and Evaluation Capacity Building for Program Improvement Field Guide, Version 1*. U.S. Centers for Disease Control and Prevention, Atlanta, GA. The CDC cannot be held responsible for material that is adapted but not directly reproduced from the above guide.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

LIST OF ABBREVIATIONS

AIDS   Acquired Immunodeficiency Syndrome
ANC   Antenatal Care
ARV   Anti-Retroviral Therapy
BC   Behavior Change
CAP   Country Assistance Plan (GAP-specific planning document)
CDC   Centers for Disease Control and Prevention
C&T   Care and Treatment
DHS   Demographic and Health Survey
FHI   Family Health International
GAP   Global AIDS Program
GFATM   Global Fund to Fight AIDS, Tuberculosis and Malaria
HHS   Health and Human Services
HRSA   Health Resources and Services Administration
HIV   Human Immunodeficiency Virus
HMIS   Health Management Information System
HQ   Headquarters
MAP   World Bank Multi-Country HIV/AIDS Project
M&E   Monitoring and Evaluation
MERG   Monitoring and Evaluation Reference Group of UNAIDS
MIS   Management Information Systems
MOH   Ministry of Health
NAP   National AIDS Program
NGO   Non-governmental Organization
OD   Office of the Director
OGH   Office of Global Health
OMB   Office of Management and Budget
ORC   Opinion Research Corporation
PBRS   Planning, Budgeting, and Reporting System (GAP-specific database)
PMTCT   Preventing Mother-to-Child Transmission
STIs   Sexually Transmitted Infections
TA   Technical Assistance
TB   Tuberculosis
UNAIDS   Joint United Nations Programme on HIV/AIDS
UNGASS   United Nations General Assembly Special Session on AIDS
UNICEF   United Nations Children's Fund
USAID   United States Agency for International Development
UTAP   University Technical Assistance Project
VCT   Voluntary Counseling and Testing
WB   World Bank
WHO   World Health Organization

PROLOGUE

PROLOGUE
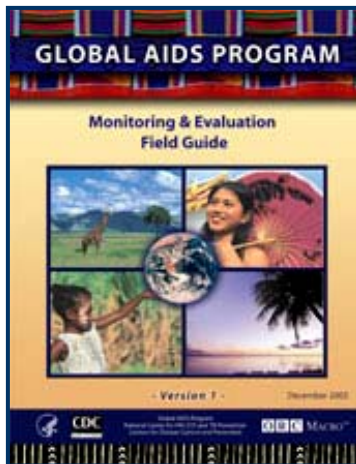
This field guide is provided as a basis for building the capacity of HIV/AIDS-related programs and initiatives throughout the world with the aim to improve programs and measure the success and impact of projects and countrywide programs. The guide addresses ways in which monitoring and evaluation data may be used to inform and improve programs implemented in partnership with the Centers for Disease Control and Prevention's (CDC's) Global AIDS Program (GAP) and other organizations working with National AIDS Programs (NAPs).

Organizations working collaboratively contribute to the overall effectiveness and impacts of HIV/AIDS interventions in each country. As such, these organizations are jointly responsible for enhancing capacity for tracking national-level trends. Therefore, GAP is committed to strengthening national HIV/AIDS-related monitoring and evaluation (M&E) systems to enhance the quality and timely use of M&E data for improving NAPs and global reporting of trends in program implementation and outcomes. National-level monitoring requires intensive data-collection efforts such as large surveys or comprehensive health information systems. Therefore, collaborative partnerships and joined resources are typically necessary to implement these M&E efforts. These partnerships require ongoing planning, negotiation, and buy-in within an agreed-on framework.

To leverage resources and sustain M&E efforts, both GAP headquarters (HQ) and field staff are encouraged to seek and develop partnerships for designing and implementing GAP-supported M&E activities at the national and sub-national levels. The nature of these partnerships might range from information gathering and information sharing to joint funding, technical inputs, and implementation. External partnerships at global and country levels, will be critical to the initiation and success of several key GAP M&E

activities, especially related to the assessment of higher-level impacts that require population-based data collection. The United States Agency for International Development (USAID) and its cooperating agencies have been and will continue to be a very significant partner in development of national M&E systems. Other major external partners include other U.S. Health and Human Services (HHS) agencies, such as the Health Resources and Services Administration (HRSA); the United Nations (UN) Agencies such as the Joint United Nations Programme on HIV/AIDS (UNAIDS), the World Health Organization (WHO), and the UN Children's Fund (UNICEF); the World Bank; and the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM).

The Multi-Agency M&E Logic Model in Figure 1 below depicts the way in which various donor agencies track outputs that are directly related to each agency's supported program/activities (e.g., a number of clients receive HIV counseling and testing). These program outputs lead to short-term and intermediate outcomes (e.g., individuals at risk of HIV develop and commit to a risk-reduction plan), which may be overlapping with expected outcomes of other agencies' work. As longer-term outcomes (e.g., behavioral change reinforcing HIV risk-reduction) and impact indicators (e.g., decrease in HIV incidence) are considered, these become more and more the result of collective contributions from all programs in each of the respective countries.

Significant groundwork has been laid for coordination of international M&E standards under the auspices of the Monitoring and Evaluation Reference Group of UNAIDS (MERG). MERG was established in 1998 to advise UNAIDS on global and national M&E, but also serves in bringing together major donors, NGOs, and technical experts in the field of M&E to form consensus on M&E standards and M&E guidance documents for use in host countries. The MERG has contributed substantially to strengthening M&E in host countries. CDC/GAP M&E is an active member of the MERG.

Figure 1. Multi-agency M&E Logic Model
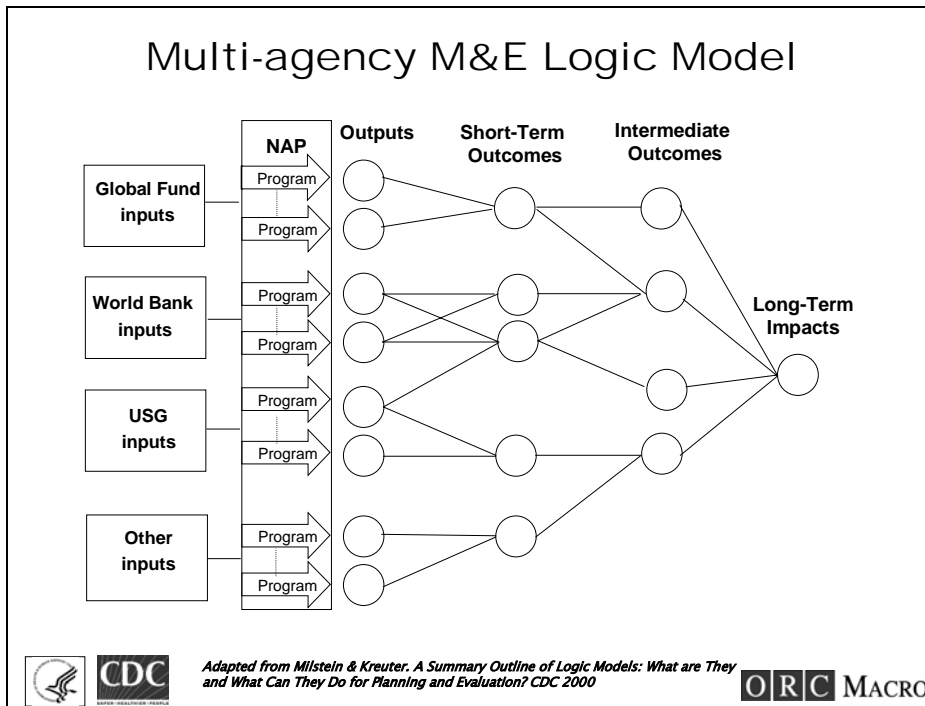
Methods of Data Collection and Data Systems

A variety of data sources and systems are important in the measurement and management of national and global indicators. Principally, these include:

- Administrative records and national reports.
- Facility-based management information systems (MIS).
- Facility surveys (service provision assessment).
- Population-based surveys.
- Disease surveillance and systems.
- Synthesis, estimation, data flows, and national databases.

Each of these will be briefly described.

1. Administrative Records and National Reports

Under this heading are included all of the various sources of information that would be used to describe program inputs and program-related and project-level activities, short of special community and facility surveys, and routine service-level data (MIS), which are described afterwards. These include:

- Budgets and Expenditure records.
- Logs of commodities, e.g., drug and test kit, receipt and distribution.
- Reports of NGOs/CBOs that include counts of program inputs and activities.
- National reports on strategic frameworks and plans.

2. Facility-Based MIS

Because of the growing emphasis of GAP programs on facility-based interventions such as preventing mother-to-child transmission (PMTCT), care/anti-retroviral therapy (ARV), and voluntary counseling and testing (VCT), investments in facility-based MIS are necessary to allow assessment of service uptake and coverage. As programs move from demonstration in nature to national (expanded) programs, decisions regarding the nature of these investments will need to be made. Field offices are encouraged to assist in development of a national MIS, e.g., CDC-Thailand for PMTCT. To promote a unified national system, a consensus on forms and software applications will need to be developed amongst key national players (e.g., Government, other bilateral agencies and international groups, and private sector).

More than any other data collection approach, the success of a MIS depends on its integration within program implementation. This presents challenges with regard to implementation, but great opportunities to generate data at a locus where it can be used directly for service improvement. GAP-informatics is currently working with partners to expand capabilities to provide direct support for national health-related MIS.
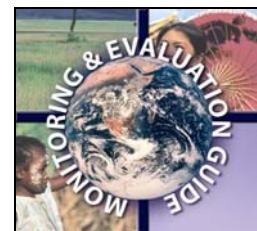
Figure 2 provides a demonstration of data results from BOTUSA-supported MIS for VCT services, showing expansion from 2000 to 2003 in Botswana's Tebelopele Network of Centers.

Figure 2. Example of National Data Results, Botswana

Monthly Reports of VCT Service Use at Tebelopele VCT Center Network sites Botswana, Apr 2000 – March 2003 (total users = 60,666); source: BOTUSA
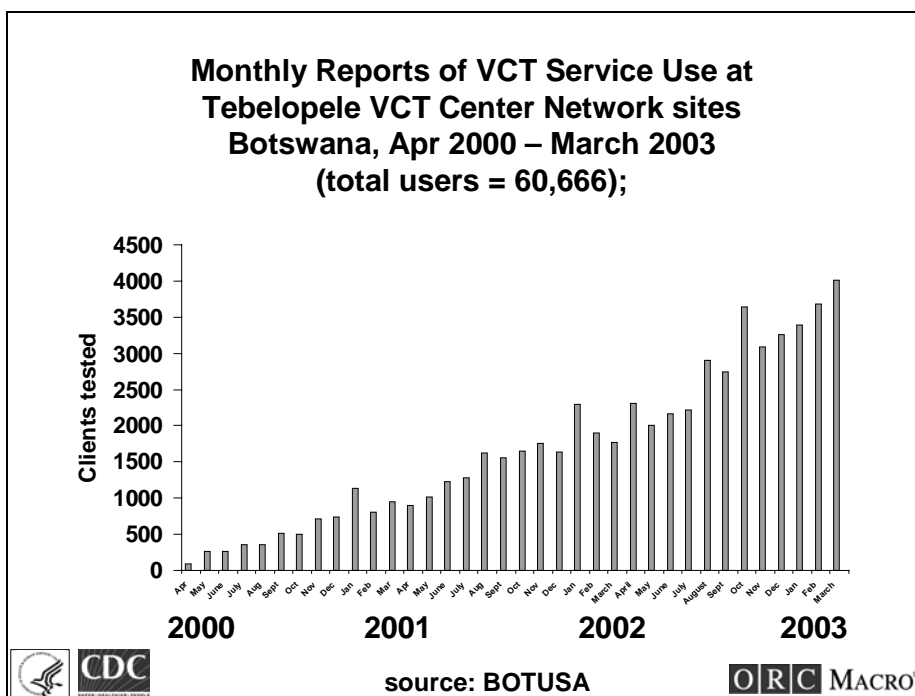
3. *Facility Surveys (Service Provision Assessments)*

In its most basic form, a facility survey involves a site inventory of all elements required to deliver services: basic infrastructure (i.e., private counseling rooms, running water), drugs, equipment, test kits, registers and, importantly, staff trained in the delivery of the reference services.   Because of the considerable cost involved in fielding facility surveys (or service provision assessments), they need to satisfy a broad data-use constituency (i.e., multiple objectives) and thus cover a wide range of services at sites.   The units of observation are facilities of various types and levels in the health system and will normally include both public and private facilities in the sample frame of sites.

The success of national HIV/AIDS program scale-up is in part assessed based on the percentage of potential service delivery sites with the capacity to deliver core services. The data to assess this type

of service-availability indicator will most often be lacking in the respective national MIS. For instance, the existence of trained staff and required infrastructure at the service delivery point can only be well measured by visiting facilities in a facility survey context.

HHS/CDC Uganda, in partnership with USAID, funded and technically supported the design and implementation of the 2002 Uganda Health Facility Survey. The type of survey findings highlighted below in Figure 3 are crucial to informed decision-making.

Figure 3. Example of Facility Survey Findings, Uganda

Percentage of Clinical Facilities Reporting Selected HIV/AIDS Services, According to Operating Authority. Uganda Health Facility Survey 2002

Source: MOH/Uganda
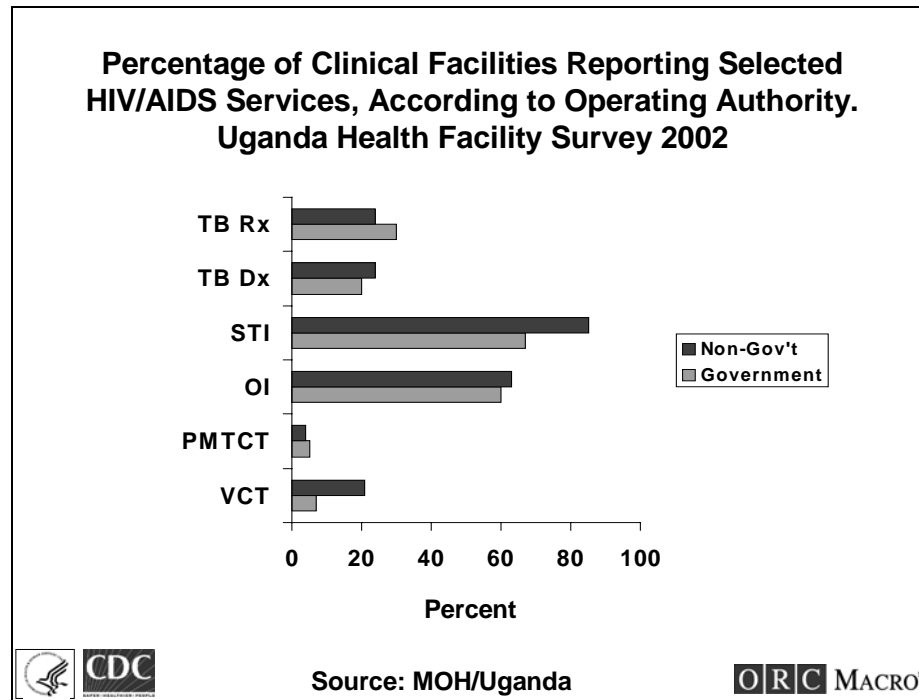
Facility surveys are typically extended to include collection of data on provider knowledge and attitudes, and observations of the quality of service delivery.

In sum, as we move from pilot approach to expanded access to services and national coverage, GAP will need to partner with national government and other donors to plan, fund and conduct

periodic (sample) facility surveys.[1] This is especially true as involvement broadens in a wide range of facility-based HIV/AIDS interventions (e.g., care, counseling and testing, blood safety, etc.).

4. *Population-based Surveys*

National population-based surveys, like the Demographic and Health Survey (DHS), are a critical component of an overall national M&E plan. Data that can only be collected through interview with individuals include: sexual behaviors, knowledge of health services, and HIV/AIDS-related stigma (in the population). For instance, a community's preparedness to use PMTCT services depends in part on awareness of mother-to-child transmission and knowledge and acceptability of the intervention components: HIV-testing at antenatal care (ANC) and delivery sites, ARV prophylaxis, and primary prevention and infant feeding counseling approaches. Efforts to increase the supply of new services need to be balanced with community mobilization and demand creation activities, which are assessed with community or population-based surveys.

GAP field offices may choose to invest in collection of data on HIV/AIDS-related knowledge, attitudes and behavior within the context of large-scale national health surveys, like a DHS (AIDS module). The disadvantage of this approach is that such broad-based surveys are implemented only every 3-5 years and do not typically provide scope for significant country-level adaptation. A more strategic investment may be made in collaboration with country partners on a national AIDS Indicator Survey, models of which are being developed by USAID and CDC in partnership with UNAIDS and other international groups. This allows the right balance between standardized instrumentation and country-specific modification. As well, a national survey apparatus such as this provides a unique opportunity for piggybacking of biomarker data collection (e.g., HIV) to produce nationally representative estimates to complement sentinel surveillance data. The 2001 Young Adult Survey, supported by GAP Zimbabwe, is a successful example of this approach.

---

[1] During early scale-up period, use of simple facility line lists could assist is making gross estimates of progress in expansion of site availability of services.

Smaller-scale community surveys, using simple instruments and limited sample designs, can also be considered to make baseline and then periodic assessments of catchment population "preparedness" for uptake of services being newly introduced.

5. *Disease Surveillance Systems*

Measurement of the impact of HIV/AIDS programs lies ultimately in tracking of changes in infection, disease, and death rates. CDC/GAP is well invested in the support of national efforts to track HIV seroprevalence, largely through sentinel surveys of antenatal clinic samples. As mentioned above, increasing interest involves use of population-based approaches. However, HIV seroprevalence is a rather blunt instrument for developing understanding of both epidemics and responses to epidemics.

The necessary complement to national seroprevalence data is in enhanced systems to collect, analyze and report on:

- Trends in high-risk behaviors (see population-based surveys, above).
- AIDS and AIDS-related disease conditions.
- Mortality trends.

With some exception, these systems are largely undeveloped or too incomplete for use in national interpretation. GAP offices may choose strategic investment in development of these information systems to provide a fuller epidemic picture and (in combination with good program data) allow assessment of the collective effectiveness or impact of national responses.

6. *Synthesis, Estimation, Data Flows, and National Databases*

Figure 4 below provides an example of a national schedule or plan for data collection and use activities for HIV/AIDS program M&E.

**Figure 4. Example of National
Schedule for HIV/AIDS Program M&E**
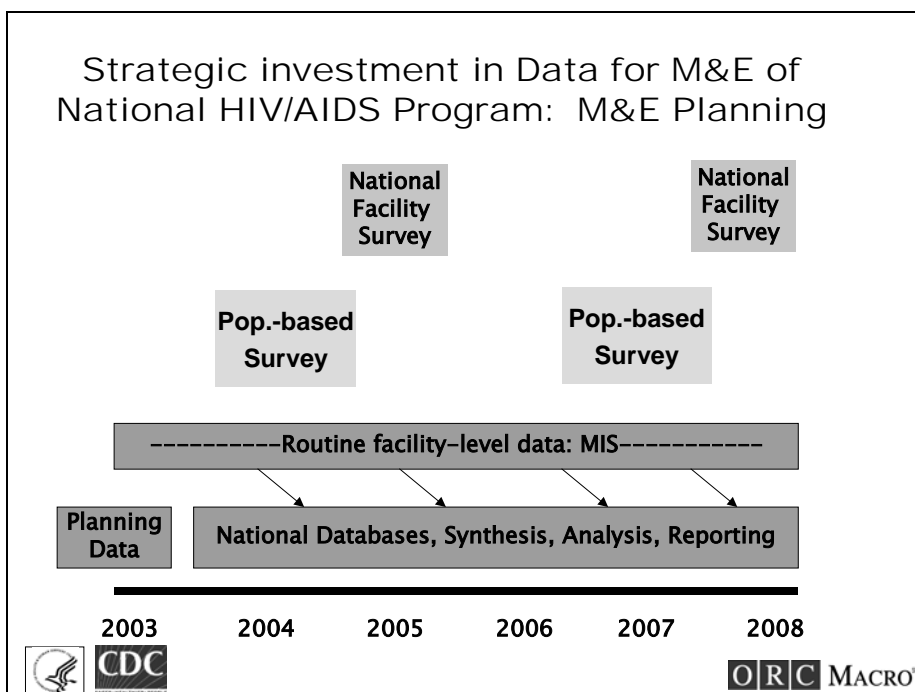
Both continuous (routine) and periodically conducted surveys and synthesis, estimation, and reporting efforts are part of a successful national M&E plan (in this example, over a 5-year period). Decisions on when and where to invest in the implementation of the national M&E plan will consider the following:

- What is your assessment of need for the data?
- Is the proposed approach the most cost-effective means?
- Is there a funding or technical gap not currently filled?

The human and other resources required to convert the various data types to useful information should not be underestimated. Support for a centralized national database (e.g., at the NAC or elsewhere) is another potentially effective means to enhance national M&E performance. This centralized unit will ideally function in the following capacities: ongoing assessment of data quality, generation of national indicators estimates, analysis, synthesis, reporting, coordination of regional/state M&E offices, and coordination with international partners.

With this overview of GAP's role in conducting M&E in conjunction with NAPs and in collaboration with other international partners, this field guide is presented as a starting point for, and overview of, conducting M&E of GAP programs. This guide is directed to GAP field and regional offices, but most of the information can be used in a generalized fashion by a number of partners and entities interested in HIV/AIDS related M&E.

PART I: INTRODUCTION AND OVERVIEW
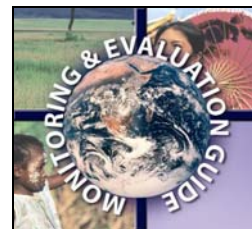
1. SETTING THE STAGE

A steadily growing HIV/AIDS epidemic worldwide and a dramatic increase in attention and funding for this epidemic have led to a global demand for accountability. Consequently, needs exist for the development of better Monitoring and Evaluation (M&E) systems. This situation creates the need for greater capacity to plan and implement M&E systems that generate information in a timely manner. This knowledge can then be used to more effectively design and improve programs, and M&E information from worldwide partners will indicate if we are collectively making a difference.

This Global AIDS Program (GAP) monitoring and evaluation (M&E) field guide was written for use by GAP field offices. It was designed to provide guidance in planning, implementing, evaluating, and improving GAP-related activities and programs. Whether readers have extensive or little experience in M&E, this guide contains information and tools needed by program staff to demonstrate program effectiveness and provide other important information for the public, funders, and peers engaged in similar work. As such, this guide offers information about necessary M&E approaches for program improvement, as well as M&E activities and related data reporting.

Many resources are available for program staff who want to learn more about M&E. The intent of this guide is not to merely repackage these resources, but to connect basic M&E principles with specific GAP-related activities, using CDC's framework for monitoring and evaluating programs. To accomplish this goal, chapters in this guide contain:
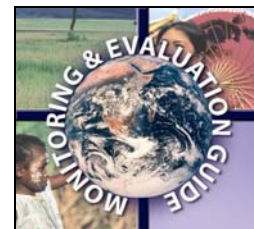
- GAP-specific and general M&E definitions and conceptual understandings.

- Worksheets and suggested activities that field office staff and partners may use as tools to assist in compiling M&E Plans and an integrated M&E strategy.

- Suggested written and Web-based resources and M&E reference materials.

Part I, Chapter 1 sets the stage for the guide, providing basic background information. An overview of the GAP approach begins in Chapter 2 with a brief description of GAP and its intended outcomes. This explanation is followed by a discussion in Chapter 3 of the basics of M&E, providing definitions of key terms and descriptions of the various types of M&E activities. These and other key M&E terms also appear in the Glossary in Appendix A. Synonyms of key terms and corresponding terms in French also appear in this Glossary.

Part II addresses the use of M&E for program improvement. Chapter 4, on developing program goals and related measurable objectives, describes a fundamental component of program development that precedes and is inextricably related to M&E. The development of program goals and objectives is critical in meeting GAP expectations. For instance, the development of a Country Assistance Plan (CAP), (which is further addressed in Part III), requires field offices to write clear, effective goals and objectives. But, beyond meeting the requirements in developing the CAP, the work of writing clear goals and objectives is an important early step in designing effective programs.

Chapter 5, Describing the Program, relates to building a logical framework for a single program or multiple programs. The use of a "logic model" provides a description of programs, referencing both the specific problem or situation that the program is designed to address and the objectives that it attempts to achieve. In addition to assisting program staff in describing and planning a program, logic models permit planners to conceptualize an evaluation approach that focuses on assessing achievement of various program stages identified by the logic model. Chapter 6 answers the question of who needs information about program implementation—in other words, who are the users of M&E data and how will they use the data. It is important to ask this question of use and users prior to planning and designing data collection strategies—how data will be used should, in fact, drive what data are actually gathered.

Chapter 7 describes ways in which the above identified data users can then assist in developing M&E questions; Chapter 8 provides a discussion on sources of data to answer various types of M&E questions; and Chapter 9 addresses approaches and systems for storing, managing, and analyzing these data. Chapter 10 discusses the dissemination of findings and uses of M&E data. Chapter 11 helps program staff assess their capacity to conduct M&E activities and offers recommendations for enhancing this capacity, including guidelines for identifying and working with outside evaluators and technical assistance (TA) providers.

Part III addresses GAP Planning and Reporting Requirements. Chapters 12 through 14 provide an overview of the GAP M&E Framework and Expectations; GAP's Planning, Budget, and Reporting System (including templates and worksheets in Appendices that provide guidance in required documentation and reporting); and an overview of the M&E Plan, which is the document required of all field offices to outline planned and ongoing M&E activities. As stated in the disclaimer at the beginning of this document, the GAP planning, budget, and reporting templates addressed in Chapter 13 and Appendices K, L and M set out GAP-specific planning and reporting requirements for the reporting round completed in 2003. These requirements may be updated for future reporting rounds, hence GAP offices should refer to guidance sent out from the GAP Office of the Director each year to meet their reporting requirements. The issues and templates in this guide, written prior to implementation of President Bush's Emergency Plan for AIDS Relief, are for illustrative purposes only.

After completing a review of the guide, field office staff should be able to:

1. Describe the use of M&E data for program planning, management, and improvement.

2. Understand and initiate program monitoring and evaluation activities as standard program components within technical strategies.

3. Work towards developing technical strategy-specific M&E plans and an integrated M&E strategy.

4. Identify M&E TA needs.

---

2. UNDERSTANDING THE GAP M&E CONTEXT

GAP began in 2000 to provide international assistance in:

- Preventing HIV infection.
- Improving access to HIV/AIDS care and treatment.
- Building the capacity and infrastructure of National AIDS Program staff to manage and evaluate their programs.

In collaboration with global partners, GAP achieves these ends by providing financial and technical assistance to governments and national and international entities working in resource-constrained countries. To date, GAP operates in 25 countries and five regional centers in Africa, Asia, and Latin America and the Caribbean.

GAP maintains technical competence and support in 17 technical strategies shown in Figure 5 below. These strategies collectively address surveillance, HIV prevention, AIDS care and treatment, monitoring and evaluation, and capacity and infrastructure support. As part of capacity building, GAP has developed an M&E strategy intended to support GAP programs in understanding, developing, and implementing data collection, management, analysis, utilization, and reporting. This strategy was drafted early in GAP's existence by a working group of international experts. Today, this M&E strategy is implemented by GAP Headquarters (HQ) M&E staff through an operational plan with guiding principles, objectives and actions for supporting an M&E system for all GAP country programs, activities, and regional centers.

Importance of M&E

Monitoring and evaluation are fundamental aspects of good program management on all levels; e.g., local, national, and global.

Monitoring: The routine tracking and reporting of priority information about a program and its intended outputs and outcomes.

Evaluation: A rigorous, scientifically based collection of information about program activities, characteristics, and outcomes that determine the merit or worth of a specific program. Evaluation studies are used to improve programs and inform decisions about future resource allocations.

M&E:

- Provides data on program progress and effectiveness.
- Improves program management and decision-making.
- Allows accountability to stakeholders, including funders.
- Provides data to plan future resource needs.
- Provides data useful for policy-making and advocacy.

Figure 5. CDC/GAP Program Model
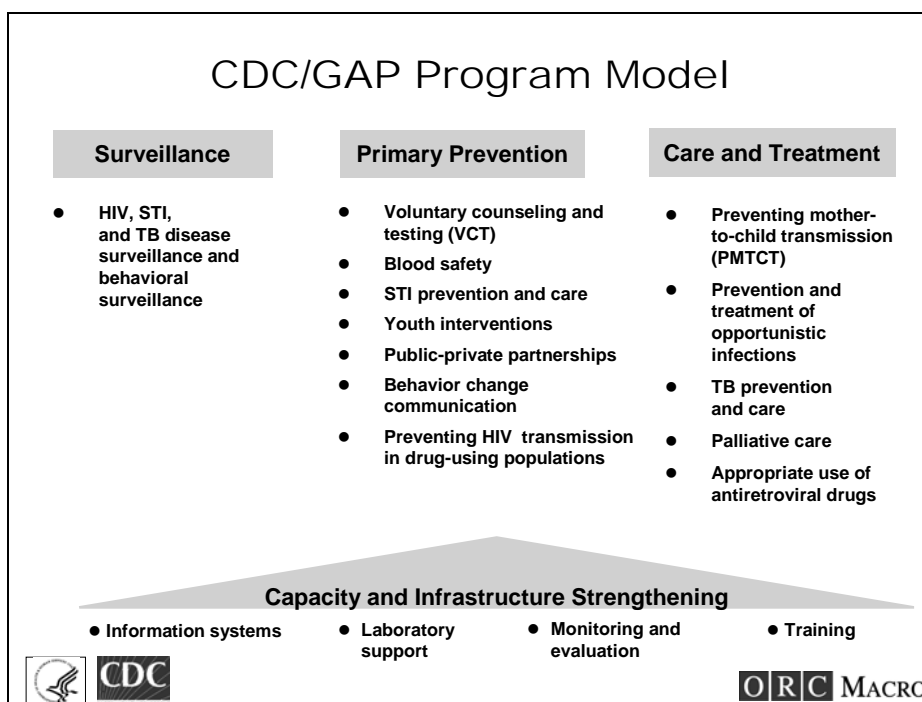
Because of the significance of conducting M&E activities, M&E Plans for various technical strategies are required of GAP field offices. GAP has devoted attention and resources to M&E from the onset of the program and will continue to invest approximately 10% of program resources to M&E.

M&E Mission and Goals

Implementation of M&E is guided by the mission statement provided in the text box below. This mission defines the overall direction of the team's activities.

CDC/GAP Monitoring and Evaluation Mission Statement

Strengthen M&E systems to inform HIV/AIDS policy and program decisions at local, national, and global levels.

To achieve this mission the following goals were established:

Goal 1: Determine the progress and effectiveness of CDC/GAP programs and assistance activities.

Goal 2: Strengthen the capacity of National AIDS Programs to conduct monitoring and evaluation.

To achieve these goals, GAP collaborates with a number of partners, including Ministries of Health and National AIDS Programs (NAPs) in countries where GAP field offices exist, the United States Agency for International Development (USAID); the Joint United Nations Programme on HIV/AIDS (UNAIDS); the World Health Organization (WHO); the World Bank Multi-Country HIV/AIDS Project (MAP); and other partners, such as the U.S. Health Resources and Services Administration (HRSA); the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM); and the United Nations Children's Fund (UNICEF). M&E uses a number of interdisciplinary approaches for collecting and analyzing data and cuts across all GAP technical strategies, such as Voluntary Counseling and Testing (VCT), Preventing Mother to Child Transmission (PMTCT), Palliative Care, and the like. An M&E Plan is needed for each technical strategy where there are substantial resources committed by field offices. These specific Plans are then integrated into an overall M&E strategy or operational plan to ensure a coordinated effort within the country.
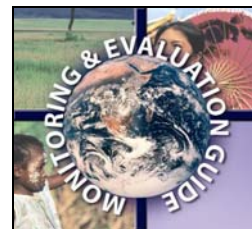
3. Understanding M&E Language

If this document were an M&E manual for a retail business, it would be fairly straightforward. The basic evaluation question likely would be: Did we make a profit during the last quarter?

If the answer to this question was "yes," executives would be pleased and would then ask, "Now, what can we do to increase the amount of profit?" If the answer was "no," these executives would be concerned about identifying the operational areas that needed to be adjusted, such as marketing, employee training, and the quality and quantity of products.

Monitoring and evaluating GAP programs is not as straightforward, but it does follow a similar logic. In the field of public health, program evaluation can tell us whether programs and initiatives are making a difference and for whom. It can identify aspects of a program that need to be adjusted or replaced, as well as areas that are on target. Information gained from M&E can lead to better decisions about program investments. Additionally, this information can demonstrate to funders and other stakeholders that their investments of time, energy, and money are providing a return.

To design M&E activities that track and assess the above program elements, it is important to understand the different types of activities and what questions they can answer. A number of different terms exist to describe M&E, and it can be confusing when different terms are used for the same type of activity. The following standard terms were adopted by CDC to discuss M&E [for further explanations, see CDC (1999) document referenced in the Further Tools and Reading section below, as well as a summary of this framework in Appendix B.]

Assessment and Planning: Assessment and planning involves the collection of information and data needed to plan programs and initiatives. These data may describe the needs of a population and the factors that put people at risk, as well as the context, program response, and resources available (financial and human). Assessment

and planning activities answer questions such as, "What are the needs of the population to be reached by the program/initiative?" How should the program/initiative be designed or modified to address the population needs?" "What would be the best way to deliver this program/initiative?" Activities such as "needs assessments" and "situational analysis" may occur as part of what is sometimes called "formative evaluation."

Tracking and Assessing Basic Program Elements

Monitoring and evaluation of programs entails tracking and assessing basic program elements, including program inputs, activities, outputs, outcomes, and impacts. These elements are defined as follows:

- *Inputs* refer to the resources used in a program. They include monetary and personnel resources that may come from a variety of sources, as well as curricula and materials.

- *Activities* are the program proceedings/actions, such as counseling sessions, material distribution, workshops, training, outreach, or specific technical assistance, that alone or in conjunction with other activities will have identifiable outputs.

- *Outputs* are the results of program activities. Outputs relate to the direct products or deliverables of the program activities, such as number of counseling sessions completed, number of people reached, and number of materials distributed. These early products of work often serve as documentation of progress for funders and other stakeholders.

- *Outcomes* are effects of the programs on target audiences or populations. These outcomes include effects of the program activities that may focus on the knowledge, attitudes, beliefs, skills, and behaviors of the target audience. Depending on the nature of the program or interventions, outcomes may also affect policies and environmental conditions affecting HIV risks and treatment.

- *Impacts* refer to the longer-range, cumulative effects of programs over time, such as change in HIV infection, morbidity, and mortality. Impacts are rarely, if ever, attributable to a single program; yet, a program may, with other programs, contribute to impacts on a population.

Input/Output Monitoring: Input and output monitoring involve the basic tracking and reporting of information about program inputs, or resources that go into a program (such as funding, number of prevention and education materials acquired for distribution,
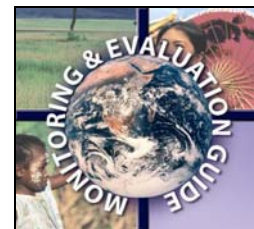
pharmaceuticals acquired for treating opportunistic infections, and staff assigned to implement the program), and about program outputs, or results of a program's activities (such as staff are trained and at-risk clients are educated about HIV-risks). This type of monitoring answers questions such as, "What services were delivered?" and "What population was served?" Data sources for monitoring inputs and outputs usually exist naturally in program documentation, such as activity reports and logs, and client records, which offer details about the time, place, number and extent of services delivered, as well as the types of clients receiving services.

Process Evaluation: This type of evaluation focuses on program implementation, adding a dimension to the information that was tracked during input/output monitoring. For example, input/output monitoring might show the number of pregnant women receiving antenatal care (ANC) sessions per quarter; the number of these women to be counseled to seek VCT per quarter; the number actually receiving HIV testing; and so on through the number of women actually receiving anti-retroviral (ARV) prophylaxis per quarter. In this example, process evaluation would add insight about these indicators by providing information about the process of implementing a PMTCT program. For instance, a process evaluation might shed light on the content and optimal time that should be allotted for sessions that encourage women to be tested for HIV; barriers to HIV testing as perceived by the clients and counselors; and perceptions that HIV positive women have about ARV prophylaxis or infant feeding practices.

Process evaluations usually focus on a single program and use largely qualitative methods to describe program activities and perceptions, especially during the developmental stages and early implementation of the program. These methods include, but are not limited to, observation, interviews and/or focus groups (with program staff, clients, and other key informants), and program document reviews. These assessments may also include some quantitative approaches, such as surveys about client satisfaction and perceptions about needs and services. In addition, a process evaluation might provide

understanding about a program's cultural, socio-political, legal, and economic contexts that affect programs.

Outcome Monitoring: Outcome monitoring is the basic tracking of variables that have been adopted as measures or "indicators" of the desired program outcomes. With National AIDS programs, outcome monitoring is typically conducted through population-based surveys to track whether or not desired outcomes have been reached. This type of monitoring may also track information directly related to program clients, such as change in knowledge, attitudes, beliefs, skills, behaviors, access to services, policies, and environmental conditions. Outcome monitoring answers questions like, "Did the expected outcomes occur?" and "To what extent did these outcomes occur?"

Outcome Evaluation: Outcome evaluations are concerned with determining if, and by how much program activities or services achieved their intended outcomes. Whereas outcome monitoring is helpful and necessary in knowing whether or not outcomes were attained, outcome evaluation attempts to attribute observed change to the intervention tested; describe the extent of program outcomes; and indicate what might happen in the absence of the program. Outcome evaluations are methodologically rigorous and require a comparative element in its design, such as a control or comparison group.

Impact Monitoring: In the field of public health, impact monitoring is usually referred to as "disease surveillance" and is concerned with the monitoring of disease prevalence or incidence. This type of monitoring collects data at the jurisdictional, regional, and national levels. All GAP programs provide data that contribute to countrywide and global disease prevalence, and many GAP field and HQ staff are involved with strengthening national surveillance systems.

Impact Evaluation: Impact evaluation deals with the rise and fall of disease incidence and prevalence as a function of HIV/AIDS programs. As such, it answers the question, "What long-term effects

Activity

The M&E Definitions Worksheet in Appendix C presents an example of a program and various data-collection activities that may be applied. Completing this worksheet may be used to "practice" matching types of data collection with the definitions of M&E activities presented in this section. It may also be used as a teaching tool for partners and constituents interested in understanding basic evaluation terminology.

does a program have on HIV infection?" Impact on entire populations seldom can be attributed to a single program or even several programs. Therefore, assessment of impact on populations usually entails a rigorous evaluation design that includes the combined effects of a number of programs on at-risk populations. Impact evaluations are resource-intensive and technically complex. Therefore, very few GAP programs will be involved in this type of evaluation.

Distinguishing Program M&E from Other Concepts

Program M&E supplements and complements other data-collection efforts, such as academic research, disease surveillance, operational research/evaluation, and policy and economic evaluations. These forms of research and evaluation are complementary but distinct.

Academic research focuses primarily on hypothesis testing in a controlled environment. It typically attempts to make statements about the relationships among specific variables under controlled circumstances at a given point in time. Evaluation is an ongoing process that incorporates many types of data with the aim of improving variables it is monitoring. In academic research, changing strategies mid-stream would be unthinkable. In program evaluation, however, with evidence that a change in service delivery would improve program outcomes, *not* changing strategies mid-stream would be unthinkable.

As discussed above in the definition of "impact monitoring," *disease surveillance* is the ongoing systematic collection, analysis, and interpretation of data to describe diseases and their transmission in populations. These data can contribute to predicting future trends and targeting needed prevention and treatment programs. Surveillance is typically performed at both the district and the national levels. Surveillance data can be augmented with more detailed information that relates to a specific program.

Operational research or *operational evaluation* applies systematic research techniques to improve service delivery. This type of research

and evaluation analyzes only those factors that are under the control of program managers, such as improving the quality of services, increasing training and supervision of staff, and adding new service components. Operational research is designed to assess the accessibility, availability, quality, and sustainability of programs.

Policy evaluations are, as the name implies, assessments of application and effectiveness of policies.

Economic evaluations use applied analytic techniques to identify, measure, value, and compare the costs and outcomes of alternative interventions.

Again, these types of research and evaluation are related to, but not the same as, program evaluation.

Further Tools and Reading

American Evaluation Association Website: http://www.eval.org

> *This site has links to a multitude of evaluation-related Web sites and publications.*

CDC. (1999). Framework for Program Evaluation in Public Health. *Morbidity and Mortality Weekly Report* 48. Overview in Appendix B. Full report may be found at: http://www.cdc.gov/epo/mmwr/preview/mmwrhtml/rr4811a1.htm. *This MMWR report details CDC's framework for program evaluation. It outlines a six-step process for public health programs to use in evaluating their interventions and operations and provides a systematic way for diverse programs to effectively and systematically conduct evaluations.*

CDC. (1992). *Handbook for Evaluating HIV Education.* Atlanta: U.S. Department of Health and Human Services, Public Health Service, National Center for Chronic Disease Prevention and Health Promotion, Division of Adolescent and School Health.

Family Health International and UNAIDS. (1998). *Meeting the Behavioral Data Collection Needs of National HIV/AIDS/STD Programmes.* Arlington, VA and Geneva, Switzerland. Proceedings from a joint IMPACT/FHI/UNAIDS workshop. Co-authored by Pisana, E., Brown, T., Saidel, T., Rehle, T., and Carael, M.

Mantell, J., DiVittis, A.T., and Auerbach, M.I. (1997). *Evaluating HIV Prevention Interventions.* New York, Plenum Press.

Comprehensive discussion on evaluating HIV prevention programs, including types of evaluation, importance of needs assessment, social and political context of evaluation, selecting evaluation design, quantitative and qualitative methods, coding and inter-rater reliability, barriers to evaluation and solutions, dissemination of evaluation results, and application of theory to HIV interventions.

Mertens, T., and Carael, M. (1997) Evaluation of HIV/STD Prevention Care and Support: An Update on WHO's Approaches. *AIDS Education and Prevention* 9, 133-145.

Patton, M.Q. (1997). *Utilization Focused Evaluation: The New Century Text* (3rd ed.). Thousand Oaks: Sage.

Peersman, G., and Sikipa, G. (2002). *Strengthening the National Response to HIV/AIDS. The Role of the CDC Global AIDS Program in Zimbabwe: A Case Study.* U.S. Centers for Disease Control and Prevention, Atlanta, GA.

Pio, A., Chaulet, P. (1998). *Tuberculosis Handbook.* WHO/TB/98.253, Geneva, Switzerland.

Rugg, D., and Mills, S. (2000). Development of an integrated monitoring and evaluation plan. In Rehle, T., Saidel, T., Mills, S., Magnani, R. (Eds*.), Evaluating Programs for HIV/AIDS Prevention and Care in Developing Countries.* Family Health International, Arlington, VA.

Rugg, D., Buehler, J., Renaud, M., et al. (1999). Evaluating HIV Prevention: A Framework for National, State and Local Levels. *American Journal of Evaluation* 20, 35-56.

Thacker S.B., and Berkelman R.L. (1988). Public health surveillance in the United States. *Epidemiology Review* 10, 164-90.

University of Wisconsin Cooperative Extension Service: http://www.uwex.edu/ces/pdande

This site has a number of very useful publications produced by their staff, including Evaluating Collaboratives: Reaching the Potential and Planning a Program Evaluation. *Most documents are available as PDF files.*

UNAIDS. (2003). *Monitoring and Evaluation Modules.* Joint United Nations Programme on HIV/AIDS (UNAIDS), Geneva, Switzerland.

UNAIDS/World Bank. (2002). *National AIDS Council's (NAC's) Monitoring and Evaluation Operations Manual.* UNAIDS/World Bank, Geneva.
http://www.unaids.org/publications/documents/epidemiology/surveillance/JC808-MonEval_en.pdf

UNAIDS/WHO/USAID. (2000). *National AIDS Programmes; A Guide to Monitoring and Evaluation.* Prepared by MEASURE Evaluation Project, University of North Carolina, Chapel Hill, NC, and UNAIDS, Geneva, Switzerland. (Also see, MEASURE Evaluation. (2002*). Strengthening Monitoring And Evaluation of National AIDS Programs in the Context of the Expanded Response.* Workshop report: Dakar, Senegal.)

UNAIDS. (2000). *Tools for Evaluating HIV Voluntary Counseling and Testing.* UNAIDS, Geneva, Switzerland.

UNAIDS. (1999). *Guide to the Strategic Planning Process for a National Response to HIV/AIDS.* UNAIDS, Geneva, Switzerland.

UNICEF/UNAIDS/WHO. (1999). *Local Monitoring and Evaluation of the Integrated Prevention of Mother to Child Transmission in Low-Income Countries.* UNAIDS, Geneva, Switzerland.

UNGASS. (2002). *Guidelines on Construction of Core Indicators.* United Nations General Assembly Special Session on AIDS, Geneva, Switzerland.

USAID. (2002). *Handbook of Indicators for HIV/AIDS/STI Programs.* USAID, Washington, D.C. http://www.usaid.gov

The Evaluation Center, Western Michigan University: http://www.wmich.edu/evalctr

This site has evaluation information and links to other sites. The Center's goal is "to provide national and international leadership for advancing the theory and practice of program, personnel, and student/constituent evaluation."

W.K. Kellogg Foundation Evaluation Handbook, W.K. Kellogg Foundation: http://www.wkkf.org/Knowledgebase/Pubs

This site has several evaluation publications, including their evaluation handbook, which is a widely used handbook that provides another perspective on evaluation as a relevant and useful program tool.

The World Bank Group. (2002). Evaluation Resources for Africa. CD-ROM. Washington, D.C., International Bank for Reconstruction and Development/The World Bank Group.

PART II: M&E FOR PROGRAM IMPROVEMENT

4. DEVELOPING GOALS AND OBJECTIVES

Purpose of Goals and Objectives

Well-conceived and well-written program goals and objectives are essential for anchoring programs and communicating program expectations to others. Specific program objectives also help focus the M&E activities by stating exactly what will be measured and assessed during the course of the program.

What are Goals and Objectives?

A *goal* is a broad statement of a desired, long-term outcome of the program. As such, goals express general program intentions and help guide the program's development. Each goal has a set of related, more specific objectives that, if met, will collectively permit program staff to reach the stated goal. *Objectives*, then, are statements of desired, specific, realistic, and measurable program results.

As an example of the relationship between goals and objectives, we might say that the goal of providing this guide is to equip GAP field staff and other readers with an understanding of the knowledge, skills, and tools needed to incorporate M&E activities into planning and implementing programs. This is a general statement about what readers should attain at the completion of reading this guide.

Some objectives related to this goal are:

- By the end of Chapter 1, Setting the Stage, readers will be able to articulate the logical flow of information of this guide, the format of the guide, and reasons for this flow and format.
- By the end of Chapter 2, Understanding the GAP M&E Context, readers will be able to articulate the importance of M&E with relation to the GAP mission, reasons for conducting M&E, and the GAP M&E mission statement and goals.

- By the end of Chapter 3, Understanding M&E Language, readers will be able to describe the standard terms adopted by CDC to define various M&E activities and distinguish program evaluation from other types of evaluation and research.

It is possible to define an objective for each chapter of the guide. In these above examples, objectives are not stated in terms of how they would be measured. However, by adding the phrase, "as measured by self-tests at the end of each chapter," these objectives would be measurable.

Developing SMART Objectives

Two basic program elements that are the focus of M&E—outputs and outcomes—are most closely related to program objectives. Outputs were defined as results of a program's activities (e.g., staff are trained, those at risk of HIV are educated about risk reduction); and outcomes were defined as effects of programs or interventions on target audiences or populations (e.g., behavioral changes reinforcing HIV risk reduction, better quality of life for those infected with HIV). Program objectives simply state these outputs and outcomes in measurable terms. (Also note that "impacts" are almost always related to goals rather than more specific program objectives.)

Objectives related to outputs are known as "process objectives." Objectives related to program outcomes are "outcome objectives."

For example, an output of a VCT program might be: "Clients tested for HIV receive their test results." This output can be turned into an objective by stating it as a target that can be measured during program implementation. For instance, the objective may be stated as: "By the end of the first program year, 98% of clients tested for HIV (assuming use of rapid testing) will receive their test results." Because an output is the result of a program activity—in this case, counseling and testing—it does not refer to the way in which clients actually respond to the activity. In other words, an output and its related objective say something about the accomplishment of the "process" of delivering a service or activity, not about the effect of these services or activities on

clients. As such, objectives related to outputs are known as "process objectives."

A desired outcome of this same program might be that clients, both those testing HIV-positive and HIV-negative, form personalized risk-reduction and treatment strategies.   The objective related to this outcome might be: "By the beginning of the second program year, 65% of clients receiving HIV test results will have developed and adhered to personalized risk-reduction/treatment strategies."   This objective is stated in measurable terms and, because it is related to program outcomes, is known as an "outcome objective."

A tool to determine whether or not objectives will be measurable and useful to program planning is the "SMART" approach to developing objectives.

Applying the SMART criteria to the outcome objective, "By the beginning of the second program year, 65% of clients receiving HIV test results will have developed and adhered to personalized risk-reduction/treatment strategies," we find the objective to be:

- Specific (clients will form risk-reduction/treatment strategies).

- Measurable (at least 65% of clients receiving test results will develop and adhere to these strategies).

- Appropriate (this objective fits with the overall goals of VCT activities).

- Realistic (it is hard to determine the realism of this objective in the abstract apart from knowing the details of the particular program. Program staff will want to determine if 65% is realistic, too high, or too low of an objective to be reached).

- Time-based (by the beginning of the second program year).

A SMART Objective is:

Specific: Identifies concrete events or actions that will take place; answers the question, "Does the objective clearly specify what will be accomplished?"

Measurable: Quantifies the amount of resources, activity, or change to be expended and achieved; answers the question, "Does the objective state how much is to be delivered or how much change is expected?"

Appropriate: Logically relates to the overall problem statement and desired effects of the program; answers the question, "Does the objective make sense in terms of what the program is attempting to accomplish?"

Realistic: Provides a realistic dimension that can be achieved with the available resources and plans for implementation; answers the question, "Is the objective achievable given available resources and experience?"

Time-based: Specifies a time within which the objective will be achieved; answers the question, "Does the objective specify when desired results will be achieved?"

The Process of Determining Goals and Objectives

Writing SMART objectives at the program planning stage is essential for articulating the accomplishments that may be reasonably expected in several months, a year, and beyond. These objectives also provide the basis for your program M&E. Though helpful in a number of ways, appropriate targets to be addressed by objectives may be difficult to estimate in early program planning stages. For instance, it may not be easy to determine the number of clients who will be reached and how much change may be expected when making these estimates months in advance of implementation. Moreover, if estimates are low and funding decisions are made based on these estimates, the program may have attracted too little funding to meet the actual demands once the program is running. On the other hand, if estimates are high and the program does not achieve its objectives, it will seem that the program underachieved what it promised to do.

A checklist of resources to help estimate objectives includes:

- Your own or others' past experience with the program.
- Unpublished evaluation data (e.g., input/output monitoring data).
- Published evaluation studies.
- Discussions with experts.
- Cost per client relative to available resources.

No easy answers exist for making good estimates, but, it does help to draw from previous experience as well as the experience of others to arrive at reasonable targets.

It should also be noted that many field and regional offices do not provide direct implementation of public health programs, but, instead, provide technical assistance towards the support of such implementation, funding for implementation, assistance in scaling up programs, and other supportive services. In these cases, goals may be stated in terms of the ultimate impact that field or regional offices expect to have on clients, but specific objectives should be stated in terms of the activities, outputs, and outcomes of the contribution of the respective field/regional office.

For example, if the extent of a field office's contribution was to provide a related Ministry of Health with technical assistance in start up of VCT services and staff training in VCT, then a SMART process objective might be, "By the end of the first program year, train five Ministry of Health staff, who will then provide a training of training (TOT) curriculum for district clinical staff. An outcome objective

might be, "By the second quarter of the second program year, 100% of district clinical staff trained to conduct VCT will meet [name of country protocol and guidelines document] standards for providing pre-test counseling, HIV testing, and post-test counseling."

Writing Objectives for Reporting Program Planning

When reporting program plans to constituents and donors, it may be too onerous a task to report each and every project objective that program staff writes. Therefore, it may be necessary to integrate a number of these more particular objectives into more concisely stated objectives. The below example illustrates this point.

As part of its Tuberculosis (TB) Prevention and Care Strategy, a GAP country field office proposed to work with the country's Ministry of Health to improve use of TB register data by developing and supporting an electronic register. GAP staff developed the following specific activities and objectives:

1. Pilot and evaluate a Window-based Electronic TB Register (ETR) in 1 province by 3/31/03.
2. Decrease the transfer-out rate to < 7% in 8 provinces within 12 months of implementation of ETR (baseline in 2000 was 11.1% transfer-out rate).
3. Decrease the interruption rate to < 10% in districts within 12 months of implementation of ETR (baseline in 2000, 15.1%).
4. Decrease missing data to < 10% within 12 months of implementing ETR.
5. Conduct TA visits in 8 out of 9 provinces and support implementation of the ETR particularly regarding validation and use of data.

Points 1 and 5 represent activities that GAP staff hoped to implement and points 2, 3, and 4 are expressed as very specific and measurable objectives. Rather than listing all these objectives in the Country Assistance Plan, which has included reports on program objectives for GAP HQ, the above were expressed as one SMART objective:

- Improve use of TB register data (as measured by the decrease in the transfer-out rate (<7%), interruption rate (<10%), and missing data (<10%), as well as improved validity of data) in 8 of 9 provinces.

Examples of Objectives from Select GAP Field Offices for FY 04

I. Guidelines Development

Country Context: CDC/GAP Cambodia serves on the antiretroviral (ARV) therapy technical working group that is developing national policies and guidelines for ARV use.

OBJECTIVE: By the end of the 2nd quarter, complete and disseminate to all districts, the final guidelines on appropriate use of ARV drugs.

Major Activities: Complete draft guidelines; hold consensus meeting; obtain MoH approval; disseminate guidelines.

Required Indicators:* CDC/GAP supported strategic planning and/or the development of policies/guidelines for provision of HIV/AIDS care and treatment services [10.1]; number and percent of districts with at least one public health facility with access to a laboratory with the capacity to monitor ARV therapy according to national or international guidelines [10.7]; number of CDC/GAP-supported facilities providing ARV treatment [10.8].

II. Program Implementation

Country Context: Guyana MOH has requested CDC/GAP and USAID assistance to expand and strengthen the Pilot PMTCT program from its current 8 sites to 32 sites within the next 18 months. Planning is underway to determine the most appropriate sites to be included. A national expansion plan is being developed. CDC/GAP and USAID will be the leaders in this area.

OBJECTIVE: By the end of FY 04, at least 20 PMTCT sites located within existing public and private antenatal facilities, will become operational in Guyana.

Major Activities: (1) Develop PMTCT plan to expand services from the current 8 sites to 20 sites. (2) Select and/or renovate 20 PMTCT sites within ANC. (2) Provide necessary equipment and supplies for PMTCT at these 20 sites. (4) Recruit and train at least 12 training-of-trainers (TOTs) in PMTCT service provision who will in turn train PMTCT workers at all service sites. (5) Develop/maintain a PMTCT database. (6) Develop PMTCT drug and monitoring system.

Required Indicators:* Existence of a PMTCT drug distribution and monitoring system in the country [8.4]; number of health workers newly trained in the provision of PMTCT services [8.5]; number of health facility sites providing the minimum package of PMTCT services [8.6]; number of CDC/GAP training activities implemented for a technical program [16.1].

III. Technical Assistance Provision

Country Context: VCT is a core strategy in Uganda's National Strategic Framework and is the cornerstone for HIV prevention and care. CDC/GAP is working with partners (MOH, AIC, AIM, TASO, Mildmay) to provide technical assistance, capacity building and financial support for VCT expansion, especially to rural areas.

OBJECTIVE: By the end of FY 04, provide technical assistance for the development of training materials and HIV counseling and testing protocols for VCT.

Major Activities: (1) Recruit a VCT technical expert for CDC/GAP Uganda. (2) Develop training materials and protocols for home based VCT services. (3) Provide technical assistance to adapt training materials and protocols for couple pre- and post-HIV test counseling. (4) Provide technical assistance for the implementation of VCT services through MOH, AIC and AIM. (5) Conduct cost analysis of scale-up of finger stick testing protocols and provide feedback to MOH. (6) Develop routine screening in clinical settings.

Required Indicators:* CDC/GAP supported the development of an operational plan for expansion of VCT services [1.1]; number of VCT sites supported by CDC/GAP [1.3]; number of information/training resources developed and/or disseminated (for VCT services) [16.3].

* Indictators from <u>Monitoring the Global AIDS Program Indicator Guide for Annual Reporting</u> <u>Version 2</u>

Further Tools and Reading

Nutrition and Physical Activity Website. *Physical Activity Evaluation Handbook,* Appendix 4: How to Write SMART Objectives. National Center for Chronic Disease Prevention and Health Promotion, available at:

http://www.cdc.gov/nccdphp/dnpa/physical/handbook/appendix4.htm

Tobacco Information and Prevention Source (TIPS) Website. *Tobacco Evaluation Manual*, Chapter 2: Describe the Program. National Center for Disease Prevention and Health Promotion, available at:

http://www.cdc.gov/tobacco/evaluation_manual/ch2.html

Centers for Disease Control and Prevention (CDC). (2003). Monitoring the Global AIDS Program Indicator Guide for Annual Reporting, Version 2. U.S. Centers for Disease Control and Prevention, Atlanta, GA.

5. DESCRIBING THE PROGRAM

Note: We use a number of GAP program examples in this chapter to illustrate the steps in building logic models. These examples were compiled with assistance from a number of GAP field office staff, yet, they do not necessarily fit with the characteristics of all field office programs. Therefore, these examples are merely presented as illustration. Field and regional offices may adjust the particular information to suit local needs and circumstances.

What are Logic Models?

An important early step to conducting M&E activities is to clearly describe the program of interest. A well-described program or intervention is easier to monitor and evaluate and facilitates using M&E data to improve the program.

Logic models go by many different names. Some of them are:

- Roadmap
- Theory of Action
- Conceptual Map
- Model of Change
- Blue Print
- Theoretical Underpinning
- Rationale
- Causal Chain
- Program Theory
- Chain of Causation
- Program Hypothesis

Logic models are invaluable program design, management, and evaluation tools that describe the main elements of a program and how these elements work together to reach a particular goal, such as prevention, care and treatment of HIV/AIDS in a specific population. As described earlier, the basic elements in describing the implementation of a program and its effects are: inputs, activities, outputs, outcomes, and impacts. A logic model graphically presents the logical progression and relationship of these elements. For instance, logic models represent the relationships between a program's activities and its intended effects, and they make explicit the assumptions about how a program will effectively address a particular problem. You can use logic models to describe an entire program, parts of a program (e.g., individual projects/interventions), or multiple related programs.

As with many aspects of monitoring and evaluation, people use a variety of terms to describe logic models and their component parts. This guide will use the terms employed by GAP, but we note some other common names used to refer to logic models in the text box to the left. Similarly, there are many different ways to construct a logic model. People use a variety of visual schematics to create logic models. For instance, flow charts, maps, and tables may all be used to portray the sequence of steps that lead to program outcomes.

Why Use Logic Models?

Logic models are intended to represent the *ideal*. They describe the intended activities and their results *if things go as planned*. As such, these models help to situate and convey the way in which a program is supposed to run and what results can be expected barring unexpected barriers and changes.

The reality of changes in funding, shifting priorities, unpredictable challenges, and other stumbling blocks can lead to actual program implementation and outcomes that are much different from what was intended. Logic models can be created or revised <u>after</u> program implementation to describe the implementation process as it actually occurred and outcomes that were achieved. Since implementation and outcomes do not always go as planned, logic models are useful program monitoring tools, facilitating comparison of planned and actual implementation and enabling assessment of why differences may have occurred. Also, since logic models identify the steps necessary to reach intended outcomes, they can illuminate important evaluation priorities.

> **Logic models are useful for everyone involved in a program—program staff, funders, and other stakeholders. They increase the likelihood that program efforts will be successful because they:**
>
> - Communicate the fundamental purpose of the program.
> - Become a reference point for everyone involved in the program.
> - Illustrate program results.
> - Serve as the basis to determine whether planned activities will lead to desired results.
> - Identify potential obstacles to program operation so that staff can address problems as soon as possible.
> - Improve program staff's expertise in planning, implementation, and evaluation.

Identifying the goals of your program and developing a logic model yield many payoffs. You may learn that the program is too ambitious or not ambitious enough, or that logical connections between goals, objectives, and activities are missing.

Logic Model Components

Components of a logic model, as shown in Figure 6, typically fall within two main sections: process and outcome. The process section describes the program resources (inputs), program activities, and the direct products of the program (outputs). If the process goes as planned, it should lead to the intended outcomes and impacts.

Figure 6. Basic Logic Model

Logic Model for HIV Prevention

Assumptions/Context

⬇

Problem Statement

⬇

Implementation

Inputs   Activities   Outputs

⬇

Outcomes

⬇

Impacts
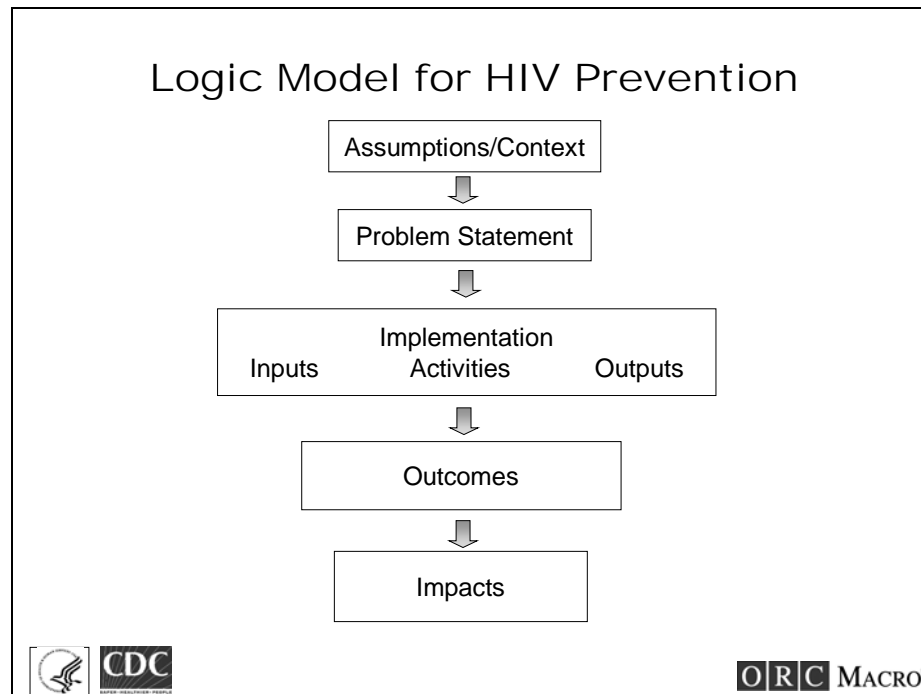
The following are explanations of logic model components:

- The **assumptions and context** relate to the unique socio-political-economic issues that exist in the locale where the respective program is being planned and the limitations and facilitators that these issues have on the potential success of the program. The assumptions that program planners make are based on the above issues, and can include theories and evidence-based knowledge that is available from similar programs. Many aspects of the assumptions and context result from assessment and planning activities. For example, a situational assessment conducted before planning a program may focus on the particular barriers and facilitators that a program will need to address to be successful.

- A **problem statement** describes the nature and extent of the problem that needs to be addressed. For example, for a population at risk of HIV/AIDS, the problem statement would include factors that contribute to the problem. These factors may be related to knowledge, attitudes, beliefs, behaviors, skills, access to services and information, policies, and environmental conditions. An example of a problem statement is:

  - HIV infection rates continue to rise, underscoring the importance for people to know their serostatus, develop personalized risk-reduction strategies, and access care and treatment services.

The problem statement often results from assessment and planning activities. For example, an assessment of prevention needs of populations at risk of HIV, or care and treatment needs of a population, may contribute to a clear and accurate problem statement.

Other logic model components include the program elements discussed in earlier chapters:

- *Inputs,* the resources used in a program, might include, for example:

  - GAP, government, and other donor funds.
  - Counseling and testing personnel (including counselors, lab technicians, and VCT site managers).
  - VCT protocols, guidelines, and training documents (Note: At the beginning of programs, inputs such as VCT protocols, guidelines, management information systems, and referral systems will need to be developed and would be considered products or "outputs" of start-up activities. When these products and systems are finally in place, they would then become "inputs" to other program elements.)
  - HIV test kits.
  - VCT data management systems.
  - Referral system for care/treatment services.

- *Activities* might include, for example:
  – Train counseling and testing personnel and site managers.
  – Provide pre-test counseling, HIV testing, and post-test counseling to all clients tested.
  – Provide pregnant HIV+ women with PMTCT services.
  – Screen HIV+ clients for (or refer to screening for) opportunistic infections and TB specifically.
  – Refer HIV+ clients to treatment services.

- *Outputs*, the results of program activities, might include:
  – Number of personnel certified in VCT.
  – Number of clients will receive pre-test counseling.
  – Number of HIV tests will be conducted.
  – Number of clients will receive test results and post-test counseling.

- *Outcomes*, the effects of the programs on target audiences or populations, might include:
  – Quality of HIV VCT increased.
  – Access to HIV testing increased.
  – Clients (HIV+ and -) formed personalized HIV risk-reduction & treatment strategy.
  – Increase in prevention, care, and treatment services for HIV+, HIV-, and discordant couples.
  – Risk behaviors decreased.
  – Health outcomes of HIV+ improved.

- *Impacts*, which are the longer-range, cumulative effects of programs, might include:
  – HIV transmission rates decreased.
  – HIV incidence decreased.
  – HIV morbidity and mortality decreased.

When discussing evaluation with others or reading other evaluation materials, a variety of different terms may be used to identify similar concepts. For instance, something identified as an output by a program may be called a short-term outcome by another program. Terms are not necessarily "correct" or "incorrect." As long as a term makes sense for a program and it is used consistently, then, the differences in terminology used by various individuals need not be a concern.

Although we have only presented components of a logic model in a list format to this point, organizations often are interested in the cause and effect relationships between activities and want to see how these elements connect. Once the more basic elements of a logic model have been described, a more detailed logic model can be developed with boxes and arrows to depict assumptions and relationships. Figure 7 shows a common format for displaying a logic model.

Figure 7. VCT Box and Arrow Logic Model
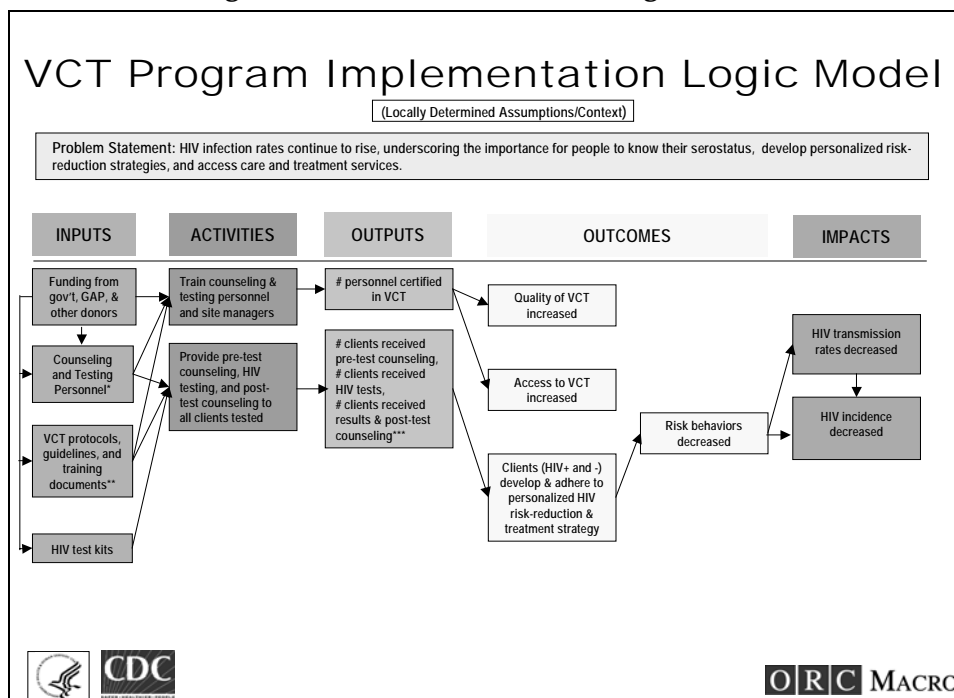
VCT Program Implementation Logic Model
(Locally Determined Assumptions/Context)

Problem Statement: HIV infection rates continue to rise, underscoring the importance for people to know their serostatus, develop personalized risk-reduction strategies, and access care and treatment services.

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES	IMPACTS

* Personnel include counselors, lab techs & VCT site managers.
** At the beginning of programs, inputs such as VCT guidelines, protocols, management information systems (MIS), and referral systems will need to be developed and would be considered "activities" and "outputs." When these products and systems are in place, they become "inputs."
*** With rapid testing, pre-test counseling, testing, and post-test counseling occur within a short time-frame. Each step is identified because it is possible that there might be a short time lag between steps that may present the possibility of some client attrition.

The above model is used to illustrate relationships between program elements, but, this model would not be sufficient to describe an entire program.

Again, logic models can be used to represent a part of a program, such as distinct projects or interventions; the complete program; or even one program in a multi-program effort. For example, two separate logic models representing VCT and PMTCT activities might be linked where activities are linked programmatically.

The above example represents the activities required to implement a VCT program; however, many GAP offices do not implement programs themselves. Instead, their funds and activities provide materials, TA, and capacity building guidance, which supports national governments in their program implementation. Logic models to describe these activities may be developed as well. Figure 8 illustrates how a logic model can be used to describe a TA program to improve the quality of VCT services. In this example, the TA program is responsible for certain inputs and TA activities, which lead to outputs. These outputs have outcomes, ultimately resulting in a greater number of sites that are able to provide high quality VCT services. In this model, the program is responsible for the outcome of improved quality of services, but additional programmatic outcomes and impacts, such as increased numbers of people receiving services, increased condom use, and decreased HIV transmission rates, are accomplished indirectly through VCT implementation by trained staff. Also, because GAP may be only one of several funding sources or TA providers, the logic model for a particular country's VCT efforts as a whole may be more elaborate than this model.

> A more complete logic model implementing a VCT program may be found in Appendix F.

Figure 8. VCT Technical Assistance Logic Model.

VCT Technical Assistance Logic Model

(Locally Determined Assumptions/Context)

Problem Statement: VCT is a critical entry point to HIV/AIDS prevention, care, and treatment services and needs to be provided consistently in a quality manner.

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES	IMPACTS

INPUTS
- Funding
- MOH & NGO staff
- VCT training curriculum
- VCT MIS
- VCT Trainers and TA providers

ACTIVITIES
- Conduct VCT Training of Trainers (TOT)
- Provide TA to trainers to train district level staff
- Provide TA to clinics to recruit & train VCT staff

OUTPUTS
- # Participants complete TOT
- # Trainers train district level clinic staff
- # Clinic staff recruit & train VCT staff

OUTCOMES
- Knowledge & skills to plan and implement VCT increased
- Knowledge of necessary referral linkages increased
- Quality of VCT increased
- # of effective VCT sites increased
- # receiving VCT services increased*
- Knowledge of HIV status increased*
- Knowledge about prevention, care, support, and Tx resources incrsd.*
- Condom use increased*
- Risk behaviors decreased*

IMPACTS
- HIV transmission rates decreased*
- HIV incidence decreased *
- HIV morbidity & mortality decreased *

* Outcomes & impacts accomplished indirectly through implementation of VCT by trained staff.

ORC MACRO

Since few programs depend on a single provider or developer, it is likely realistic to compile a "multi-agency" logic model that represents the various contributions that several entities will be expected to make to achieve the desired effects and long-term impacts on people affected and infected with HIV.

Logic models are often cyclical in that an outcome from one activity can provide information that then feeds back into a previous activity. Much of the benefit of constructing program logic models comes from the iterative process of discussing, analyzing, and justifying the expected relationships and feedback loops. Therefore, even though we actually present logic models in a "box-and-arrow" format, conceptually, they are more cyclical in nature, and the actual process

of implementing a program based on logic models may be better depicted as follows:

Figure 9. Cyclical Logic Model

Cyclical Logic Model for HIV Prevention

Process Monitoring and Evaluation

Inputs

Problem Statement

Activities

Impacts

Outputs

Outcomes

Outcome Monitoring and Evaluation

CDC

ORC MACRO
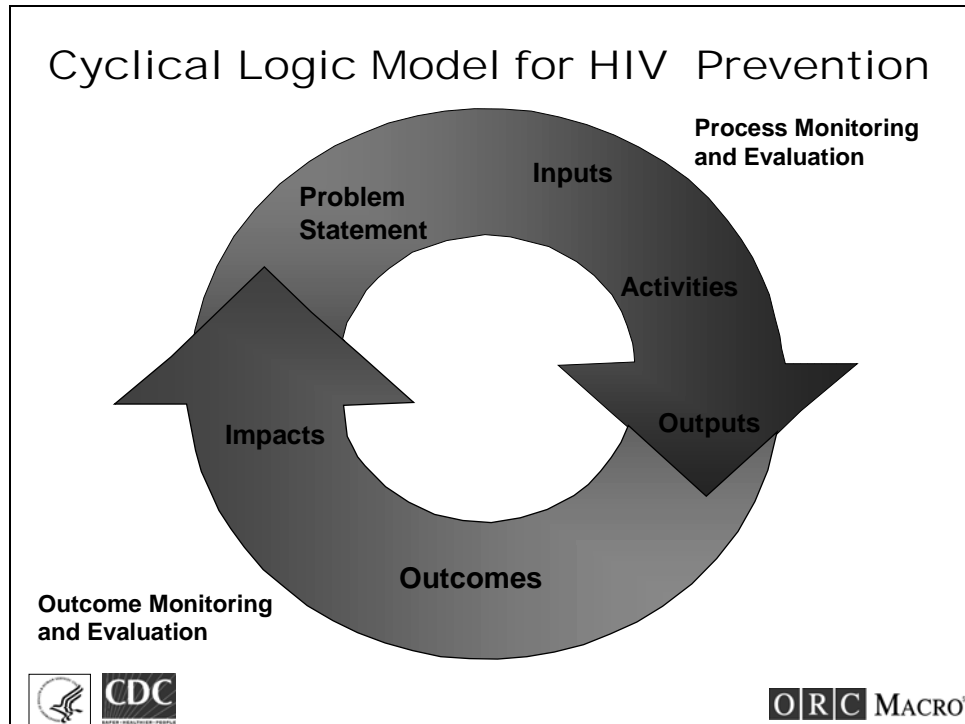
Developing Logic Models for Your Program

The VCT logic models illustrate the "big picture" expectations for the program over the long term. Depending on the availability of other resources, a logic model that describes a program may be complex, with emphasis and feedback loops in different places. In the following chapters, we also draw on sample logic models for some of the activities undertaken in GAP. If the models accurately represent a local program, they may be used as actual models that field offices implement. However, the more likely scenario is that these examples will need to be adjusted to fit local circumstances, uses, activities, and outcomes. Programs may also include activities that are not represented in these logic models. In this case it may be a very useful planning activity to engage stakeholders in creating a logic model that includes those activities.

To summarize, here are the steps for creating a logic model:

- Define problem statement.
- Specify intended outputs/outcomes, available resources, and planned activities.
- Arrange above components in a time sequence using "if, then" statements.
- Draw arrows to connect activities to intended effects.
- Review with program staff and stakeholders and refine.

Those creating new logic models may want to begin by identifying program goals and listing all the resources available for meeting them. Next, a decision may be made regarding the activities that will be necessary (and realistic, given resources) to reach these goals. As the basic model in figure 7 illustrates, logic models frequently flow from left to right, with inputs on the left, leading to activities and then, finally, on the right, to outcomes. This is not necessarily the only approach, however. Some may choose to build a model that flows from top to bottom. The key criteria are that the model is an accurate representation of a program and can be understood by stakeholders.

Ideally, the model should fit on a single page with enough detail to be explained fairly easily and understood by others. However, depicting the relationships among activities and outcomes, and uncovering the assumptions about those relationships, can be difficult. One way to proceed is to connect a chain of "If, then" statements. For example, "If we provide HIV test kits, then we will be able to test more people." This statement begs the question, "How?" The answer is another activity: conducting HIV tests. Through this process of refining the model, activities and their related anticipated outcomes are gradually identified.

Even the best logic models may need periodic revision.

As a logic model is developed, it is important to identify possible problems and solutions, as well as possible unintended outcomes. This sort of preparation—looking at the problem from all sides and imagining all the possible scenarios—will make the work proceed much more smoothly.

Logic models are useful to convey both the overview of a program and the details of program activities. An overview model, comparable to the VCT models, lays out the chain of activities and effects intended to achieve the outcomes. It clarifies intended outcomes and the range of actors to be mobilized. It can be used as a template for all of the actors as well as a template for smaller scale activities.

Working from the overview model, the focus can then be narrowed to specific activities contained in the logic model and more specific, individual logic models may be designed for each activity. These narrower models describe the activities represented by the arrows between boxes in the overview logic model. For example, you may want to set out how to move from developing local-level partnerships for HIV prevention to developing a local plan for HIV prevention; these require two separate approaches with distinct logic models. A logic model focused solely on this process will help you enumerate all the details necessary to get from one output to another.

Once a logic model is developed to provide a thorough overview of a program and its elements, a good basis will have been developed for understanding what needs to be monitored and evaluated as your program is implemented. The logic model should then naturally assist in developing M&E questions, as well as designing a plan for collecting data and measuring program progress and outcomes.

Activity

Appendix G contains a worksheet that may be helpful in describing the elements of the logic model as they pertain to particular programs.

Further Tools and Reading

CSAP Decision Support System: http://www.preventiondss.org
    Logic modeling is presented as a circular (recursive) seven-step process beginning at Assess Needs and progressing through Develop Capacity, Select Programs, Implement Programs, Evaluate Programs, Report Programs, and Get Technical Assistance and Training.

Milstein, B, and Kreuter, M. (2000). *A Summary of Logic Models: What Are They and What Can They Do for Planning and Evaluation?* CDC Evaluation Working Group. National Center for Disease Prevention and Health Promotion, CDC.

University of Ottawa Program Evaluation Toolkit:
    http://www.uottawa.ca/academic/med/epid/toolkit.htm
    Includes description of logic models, examples, and tips.

University of Wisconsin Cooperative Extension Service:

   http://www.uwex.edu/ces/pdande

   (Select *The Logic Model: A Program Performance Framework*)

   Explanation of logic models in a presentation format.

Urban Institute:

   http://www.urban.org/crime/evalguide.html#chap2

   Chapter that describes the development and uses of a logic model.

6. CONSIDERING DATA USE AND USERS

Once a program logic is articulated, ways in which programs will be monitored and evaluated may be addressed. Before developing and finalizing M&E questions that will guide data-collection needs and activities, it would be a good time for program staff to engage in a brainstorming activity to make certain that all stakeholders have been considered in plans for conducting M&E. Program staff might want to generate a list of all possible stakeholders since some of these individuals and entities may have changed since initial program planning stages.

Once an exhaustive list of stakeholders is generated, program staff may continue to be engaged by asking who among these stakeholders are the most likely to use information collected through M&E activities and how, exactly, might they use this information.

At this stage of planning for M&E, it may be good for program staff and stakeholders to think about the way data that may be generated from M&E may be used. Staff might consider both barriers to using evaluation findings, strategies for promoting the use of data, and various ways different individuals and agencies may make use of data.

Program stakeholders who may be interested in findings from M&E may include:

- National AIDS program managers and staff, as well as ministries and departments of health.
- Program clients or potential clients.
- Staff from other in-country programs.
- GAP HQ staff.
- Other funders and resource providers, such as providers of prevention materials, or care/treatment supplies.
- Policy makers.
- M&E officers/evaluators.

Possible strategies for overcoming barriers to data use include:

- Developing early buy-in from evaluation stakeholders.
- Clearly identifying the intended users of evaluation data.
- Identifying evaluation questions that are meaningful to the intended users.
- Deciding how the data will be used before the evaluation is conducted.
- After data is collected and analyzed, presenting it in a user-friendly format.

Further Tools and Reading

CDC. (1999). Framework for Program Evaluation in Public Health. *Morbidity and Mortality Weekly Report* 48. Overview in Appendix B. Full report may be found at: http://www.cdc.gov/epo/mmwr/preview/mmwrhtml/rr4811a1.htm

Mantell, J., DiVittis, A.T., and Auerbach, M.I. (1997). *Evaluating HIV Prevention Interventions*. New York, Plenum Press.

Preskill, H., and Torres, R.T. (2000). The Learning Dimension of Evaluation Use. *New Directions for Evaluation* 88.

Rugg, D., Buehler, J., Renaud, M., et al. (1999). Evaluating HIV Prevention: A Framework for National, State and Local Levels. *American Journal of Evaluation* 20, 35-56.

United Nations Population Fund. (2001). Tool No. 4, Stakeholder Participation in Monitoring and Evaluation. *Monitoring and Evaluation Toolkit for Programme Managers.* Office of Oversight and Evaluation:

http://www.unfpa.org/monitoring/toolkit/stakeholder.pdf

Available in English, French, and Spanish.

Activity

The case study in Appendix H helps in thinking about potential ways in which individuals and agencies might use evaluation findings. It addresses:

- The kinds of information or lessons that may be learned from evaluation.

- Ways in which an intervention may be improved through evaluation of its components.

- Different ways in which people use evaluation data to improve interventions or programs and implications of these for using evaluation findings.

7. WRITING AND PRIORITIZING EVALUABLE QUESTIONS

Based on the information regarding program objectives, logic models, and consideration of stakeholders, program staff should then establish the precise M&E questions that you are going to ask of your program. These questions will then determine what M&E data will be needed, as well as the necessary data-collection methods.

Writing Evaluable Questions

M&E questions are typically based on stated program objectives. Because these objectives are presented in measurable terms, it should be relatively easy to turn these into evaluable questions. For instance, a process objective might be:

By the end of the first program year, 98% of clients tested for HIV will receive test results.

The corresponding "evaluable" question is:

By the end of the year, did 98% of clients who were tested for HIV receive test results?

This would be an input/output monitoring question that would require the tracking of the number and percent of clients who were tested and received test results. This information could likely be gained by reviewing program documentation, such as client records that show whether or not clients who were tested received results.

A related outcome objective may be stated as:

By the beginning of the second program year, 65% of clients receiving HIV test results will have formed personalized risk-reduction/treatment strategies.

The related "evaluable" question would be:

By the beginning of the second program year, did 65% of the clients who received test results form personalized risk-reduction/treatment strategies?

This would be a relatively easy question to answer if clinical documentation of personalized strategies was accessible for review by VCT program staff. In cases in which this documentation was not available, program staff might need to use other means, such as asking clinical staff to give an exact or estimated percentage of clients forming personalized strategies. The second situation would be less desirable than the first, but it might be the only reasonable alternative.

In the above example, an important M&E question may have to do with the process by which clinic staff delivered information about HIV test results. A related question may involve the way in which the clients initially responded to this information: were they ready to form personalized strategies after getting results? Did it take some time before they could form these strategies? What was the process of assisting people to form strategies? Answers to these questions would likely entail staff observation of, or reflection on, the clinical process or interviews with clinic staff and, if possible, the clients themselves, to learn their perceptions about how they felt the process went. Other questions may entail learning whether or not the desired client outcomes were achieved because of the program interventions or for other reasons outside of the program. Unlike the above monitoring questions, which, to arrive at answers, would involve clinic documentation review, answers to these types of questions would entail methods beyond document review. These and other methods will be more fully discussed in the following chapter. But the primary point is that a number of questions may be asked of the program, and it is likely that there will be more potential questions than there are program resources to support answering all inquiries. Therefore, it is important to focus on the priority information necessary for decision-making.

Once program staff and stakeholders have determined priority M&E questions, the next step is deciding on the best methods for gathering data that will answer these inquiries. This is the subject of the next chapter.

Further Tools and Reading

United Nations Population Fund. (2001). Tool No. 5, Planning and Managing an Evaluation, Part II: Defining Evaluation Questions and Measurement Standards. *Monitoring and Evaluation Toolkit for Programme Managers.* Office of Oversight and Evaluation:

http://www.unfpa.org/monitoring/toolkit/5defining.pdf

Available in English, French, and Spanish.

8. GATHERING PROGRAM IMPROVEMENT DATA

M&E questions should lead to discussion of appropriate data-collection methods. Essentially, questions that have to do with the process of implementing a program will involve input/output monitoring and process evaluation methods. Questions about program outcomes will entail outcome monitoring and outcome evaluation methods. Because only a few GAP programs may be involved with impact evaluations, and these evaluation designs are much more complex than other types of assessment, we will not cover methods for assessing impact in this guide.

M&E logic suggests that evaluations occur in a staged fashion. That is, most programs that conduct outcome evaluations have implemented some level of process evaluation prior to this more rigorous assessment approach. Also, input/output monitoring data always inform process evaluation, and outcome monitoring data are pre-requisite to outcome evaluation. Instructive papers that describe this framework and logic are Rugg et al., 1999; Rugg and Mills, 2000; and Boerma et al., 2000 (see full citations below).

The below illustration of a pipeline shows this logic and depicts the expectations of GAP field offices as they relate to various types of data. As the pipeline model shows, while all field offices and projects are expected to collect and report input/output monitoring data and most will eventually collect process evaluation data, only some will conduct outcome evaluation and few will be involved in larger impact evaluation studies.

> **To briefly reiterate:**
> - Input/output monitoring tracks information about program inputs and outputs.
> - Process evaluation focuses on program implementation.
> - Outcome monitoring tracks measures related to desired outcomes.
> - Outcome evaluation collects data about if and by how much program activities or services achieved their intended outcomes.

Figure 10. Monitoring & Evaluation Pipeline

Figure content:

**Strategic Planning for M&E:
Setting Realistic Expectations**

Monitoring and Evaluation Pipeline

# of Projects

All — Input/Output Monitoring

Most — Process Evaluation

Some — Outcome Monitoring/Evaluation

Few * — Impact Monitoring/Evaluation
* Supplemented with impact indicators from surveillance data.

Levels of Monitoring & Evaluation Effort

Adaptation of Rehle/Rugg M&E Pipeline Model, FHI 2001

CDC   ORC MACRO

Input/Output, Outcomes, and Impact Monitoring

The two types of monitoring—input/output and outcome—involve data-collection methods, typically involving the review of information collected in the natural course of program implementation. For instance, inventories of prevention/education materials and pharmaceuticals, and review of program activity logs and client records would likely provide all the information needed to monitor program inputs and outputs. Similarly, client records, including results of questionnaires or surveys that test and show results of program services, usually contain ample data to track outcomes. Therefore, effective program monitoring may be accomplished with thoughtful and thorough record-keeping, the ability to aggregate data from program documents, and strict client confidentiality guidelines around information drawn from client records. The third type of monitoring—impact monitoring—typically involves the selection of key information or variables from surveillance data systems and national surveys.

As an example, consider Figure 11 below, which is a portion of a larger VCT Implementation Logic Model. For the complete model, see Appendix F.

Figure 11. Portion of VCT Implementation Logic Model

VCT Program Implementation Logic Model

(Locally Determined Assumptions/Context)

Problem Statement: HIV infection rates continue to rise, underscoring the importance for people to know their serostatus, develop personalized risk-reduction strategies, and access care and treatment services.

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES	IMPACTS

Funding from gov't, GAP, & other donors → Train counseling & testing personnel and site managers → # personnel certified in VCT → Quality of VCT increased

Counseling and Testing Personnel*

VCT protocols, guidelines, and training documents** → Provide pre-test counseling, HIV testing, and post-test counseling to all clients tested → # clients received pre-test counseling, # clients received HIV tests, # clients received results & post-test counseling*** → Access to VCT increased

HIV test kits

Clients (HIV+ and -) develop & adhere to personalized HIV risk-reduction & treatment strategy → Risk behaviors decreased → HIV transmission rates decreased → HIV incidence decreased

CDC | ORC MACRO℠

* Personnel include counselors, lab techs & VCT site managers.
** At the beginning of programs, inputs such as VCT guidelines, protocols, management information systems (MIS), and referral systems will need to be developed and would be considered "activities" and "outputs." When these products and systems are in place, they become "inputs."
*** With rapid testing, pre-test counseling, testing, and post-test counseling occur within a short time-frame. Each step is identified because it is possible that there might be a short time lag between steps that may present the possibility of some client attrition.

If program staff established as an objective that: *By the end of the first program year, 98% of clients counseled for HIV testing would be tested for HIV,* the evaluable question would be: *Were 98% of clients who were counseled, actually tested by the end of the first program year?*

To answer this question, information related to inputs and outputs would need to be tracked. Assuming that adequate funding, counseling protocols and guidelines, and personnel (inputs) were already established for this program, the following measures could be tracked and reported:

- Number of HIV test kits acquired (input).
- Number of clients receiving counseling for HIV testing (output).
- Number and percent of clients tested (output).

Programs may have established measurable objectives related to the first two measures as "prerequisite" objectives that would have to be met before the objective of testing 98% of clients who were counseled could be accomplished. For instance, program staff would likely have stated the necessary test kits that would need to be in place to serve the expected number of clients. Similarly, they would have established a target number of individuals to receive counseling in the program year before establishing the target of testing 98% of these clients.

Data sources for these measures would all likely be project documents and records kept on a continuous basis; e.g., clinic inventories of HIV test kits and records of clients.

Most of the above measures are not part of the "required" or "core" reporting indicators for annual reporting (see Appendix M). However, the final measure is similar to Program Area Indicator 1.6: *Number of individuals (by sex) tested in VCT sites supported by CDC/GAP.* If program staff kept records of numbers of clients tested by sex, this would satisfy reporting requirements, as well as assist in monitoring for program improvement.

Monitoring outcomes is similar. Program staff may have established the objective: *By the beginning of the second program year, 65% of clients receiving HIV test results will have established personalized risk-reduction/treatment strategies (outcome).*
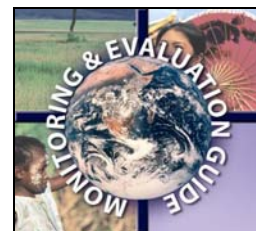
To answer the question of whether this outcome was achieved, all of the above measures would need to be tracked, as well as the measure: *Number of clients establishing personalized risk-reduction/treatment strategies.* Again, this number may be derived by project records of test results delivered to clients. Note that nothing has been said yet about the quality of VCT or any of the above steps taken to achieve the objective of having clients establish risk-reduction/treatment strategies. Neither has there been discussion about whether or not the number or percent of clients establishing these personalized strategies is an adequate outcome compared to other similar VCT programs. These are questions for process or outcome evaluations, discussed below.

Using the example above, the relationship between the VCT logic model elements, the program objective, the M&E question, and the measures that could be used to answer this question is illustrated in the following.

Relationship Between Logic Model Components and Objectives – VCT

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES	IMPACTS

| **LOGIC MODEL** | Counseling and Testing Personnel | Provide pre-test counseling, HIV testing, post-test counseling | # clients received pre-test counseling, # clients received HIV tests, # clients received results and post-test counseling | Clients (HIV+ and -) form personalized HIV risk-reduction & treatment strategy | Risk behaviors decreased | HIV incidence decreased |

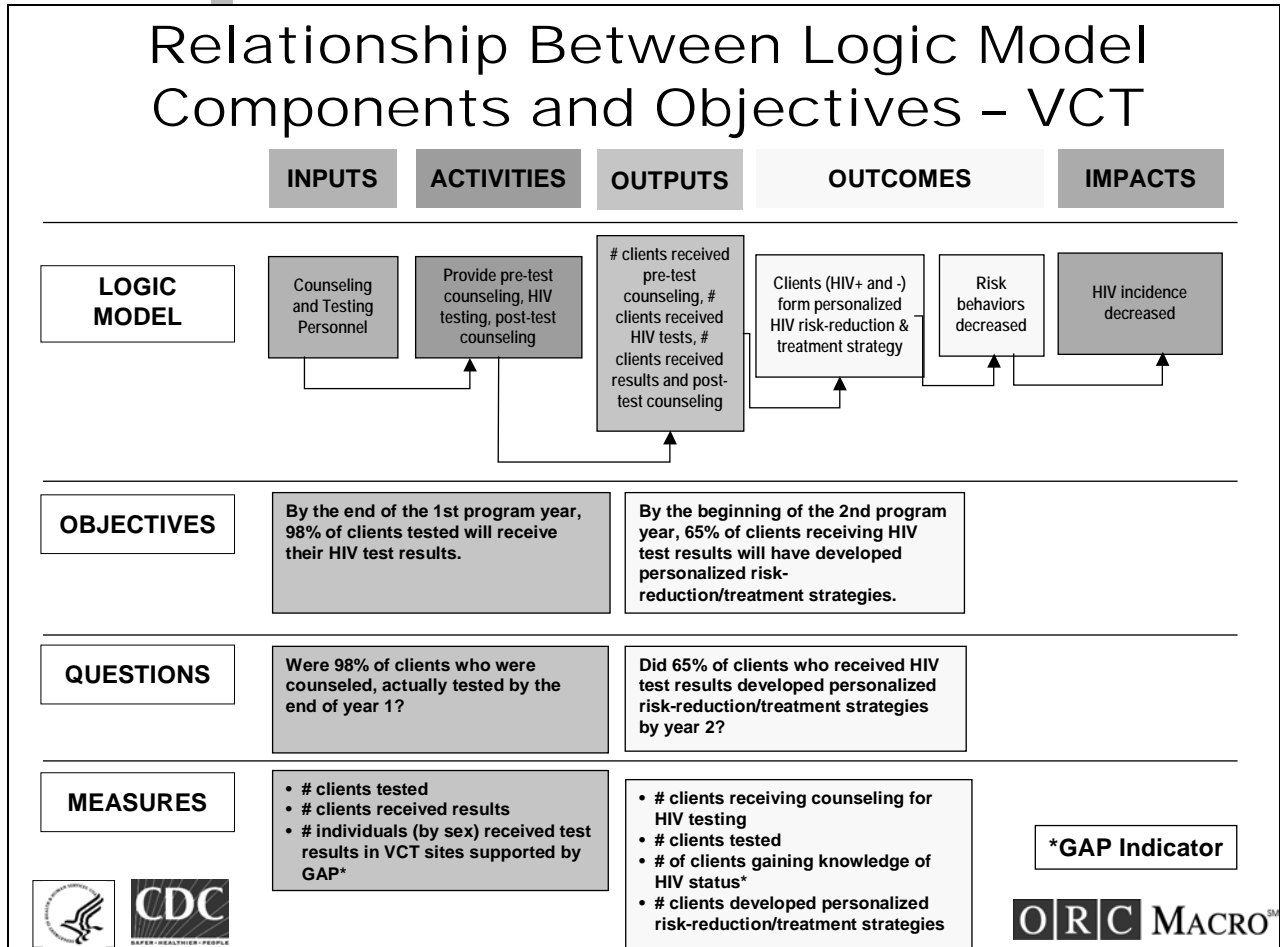| **OBJECTIVES**	By the end of the 1st program year, 98% of clients tested will receive their HIV test results.	By the beginning of the 2nd program year, 65% of clients receiving HIV test results will have developed personalized risk-reduction/treatment strategies.

QUESTIONS	Were 98% of clients who were counseled, actually tested by the end of year 1?	Did 65% of clients who received HIV test results developed personalized risk-reduction/treatment strategies by year 2?

MEASURES	• # clients tested • # clients received results • # individuals (by sex) received test results in VCT sites supported by GAP*	• # clients receiving counseling for HIV testing • # clients tested • # of clients gaining knowledge of HIV status* • # clients developed personalized risk-reduction/treatment strategies

*GAP Indicator

ORC MACRO℠

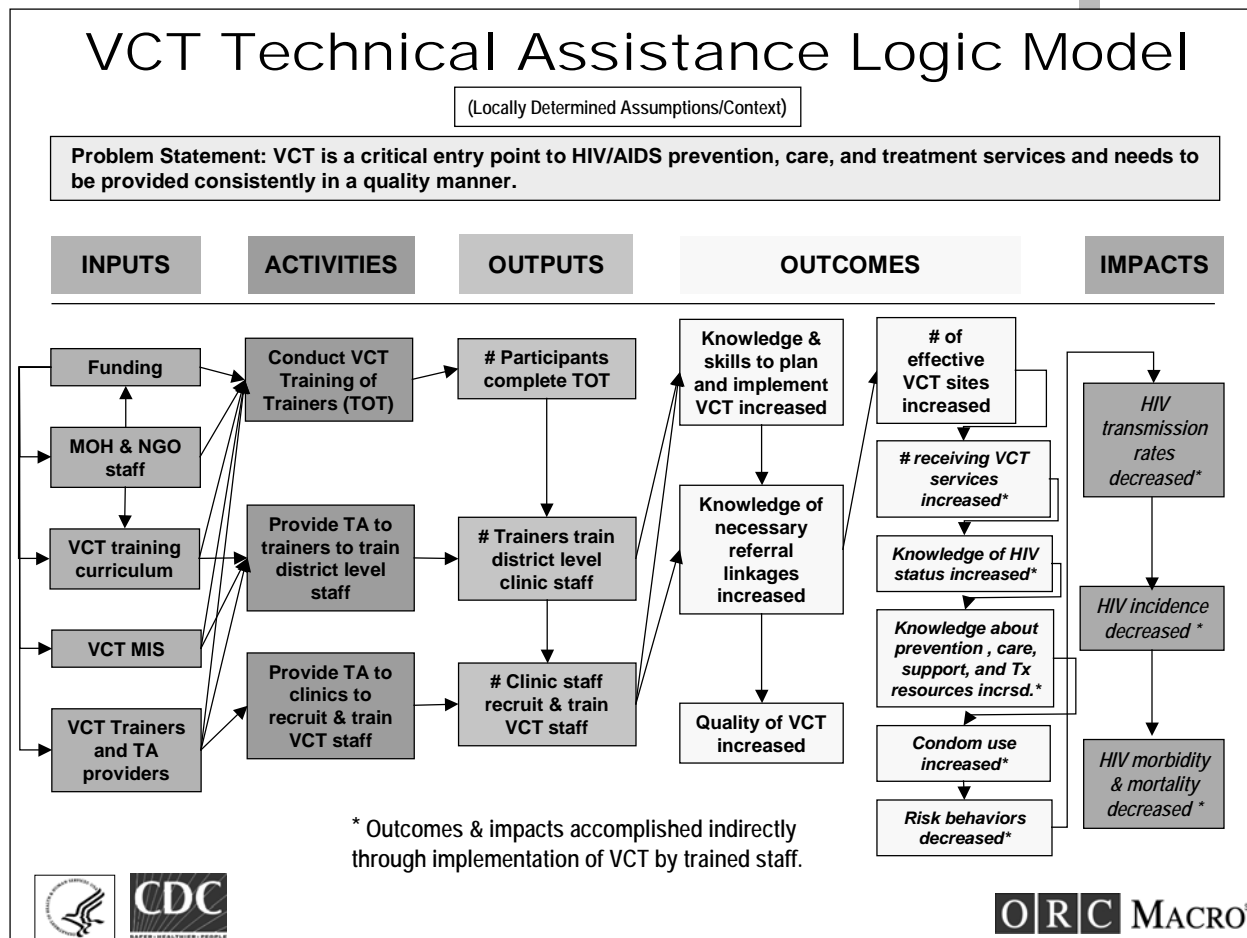In this example, it is assumed that the respective GAP VCT office supported program implementation. In some cases, however, the field office may provide technical assistance but not directly support implementation. In this case, measures will need to be established to track the provision of TA. Our earlier example of a logic model for technical assistance illustrates the point in the following:

Figure 13. VCT Technical Assistance Logic Model.

VCT Technical Assistance Logic Model

(Locally Determined Assumptions/Context)

Problem Statement: VCT is a critical entry point to HIV/AIDS prevention, care, and treatment services and needs to be provided consistently in a quality manner.

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES		IMPACTS

Funding

MOH & NGO staff

VCT training curriculum

VCT MIS

VCT Trainers and TA providers

Conduct VCT Training of Trainers (TOT)

Provide TA to trainers to train district level staff

Provide TA to clinics to recruit & train VCT staff

 # Participants complete TOT

 # Trainers train district level clinic staff

 # Clinic staff recruit & train VCT staff

Knowledge & skills to plan and implement VCT increased

Knowledge of necessary referral linkages increased

Quality of VCT increased

# of effective VCT sites increased

*# receiving VCT services increased**

*Knowledge of HIV status increased**

*Knowledge about prevention , care, support, and Tx resources incrsd.**

*Condom use increased**

*Risk behaviors decreased**

*HIV transmission rates decreased**

*HIV incidence decreased **

*HIV morbidity & mortality decreased **

* Outcomes & impacts accomplished indirectly through implementation of VCT by trained staff.

If program staff providing TA set an objective: *By the end of the first program year, 15 VCT staff were recruited and trained*, the question of whether these 15 were recruited and trained would necessitate tracking a number of measures. Assuming that funding, MOH and NGO staff, a VCT training curriculum, and a counseling protocol were established, program staff would track:

- Number of trainers trained.
- Number of district staff trained.
- Number of VCT clinic staff trained.

Data sources would be documentation of training sessions and individuals who attended these sessions. Note that the required indicators for VCT (see Appendix M) include:

- 1.2. Number of country nationals trained in provision of VCT services.
- 1.3. Number of VCT sites supported by CDC/GAP.

These indicators could easily be obtained from the information collected to track capacity-building efforts and from project documentation mentioned above. Again, using this example, the relationship between the respective logic model components, program objectives, M&E questions, and measures is as follows:

Figure 14. Relationship of Logic Models, Objectives, M&E Questions, and Measures
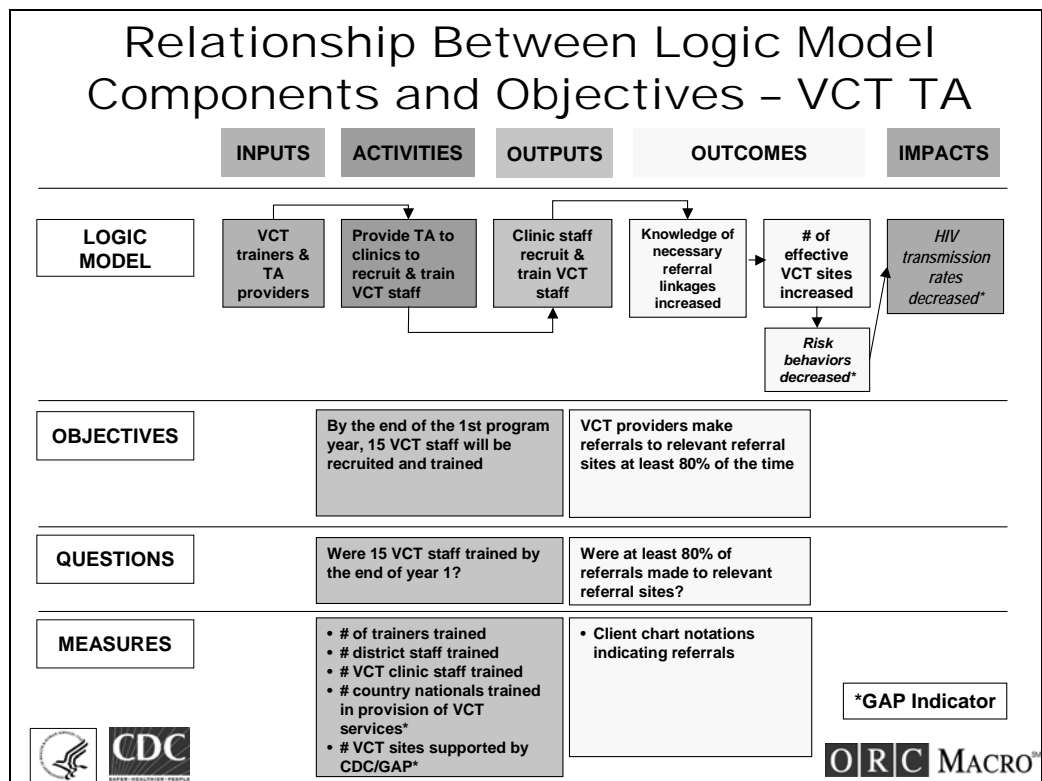
Process and Outcome Evaluation

The two types of evaluation—process and outcome—entail more detailed study designs and data-collection methods. Therefore, issues about data collection and study designs related to each will be dealt with separately below. Because of space limitations, we are only able to provide an overview of issues related to process and outcome evaluation in this guide. Each type of evaluation warrants a separate training and related manuals to complement this guide. Also, note that some evaluations may need Office of Management and Budget (OMB) and Institutional Review Board (IRB) clearances. GAP field office staff should contact the GAP Associate Director for Science (ADS) to answer questions about possible OMB and IRB review.

Process Evaluation

In general, process evaluation provides field offices and partners, service providers, and GAP HQ with information about whether the program has reached its intended audience, the level or extent and quality of services provided, and what resources were required to support the HIV/AIDS-related effort made. Contrary to some impressions, process evaluation is not something to be done when a program cannot find the resources or tools to conduct outcome evaluation. It is an ongoing activity that provides crucial information about why a program or activity did or did not work and about how to improve the program. A well-conceived and implemented process evaluation requires a reasonable amount of staff and evaluator time to collect and analyze the required information and reach conclusions.

A number of data-collection methods may be used to conduct process evaluations. Data obtained directly from program participants or staff may be collected through self-administered or interviewer-administered questionnaires and focus groups and other types of discussion groups. Data also may be obtained from program activity or participation logs and other documentation and direct observation of program implementation processes. A combination of these methods may also be employed for a fuller understanding of program processes. A more complete overview of methods is offered below.

To determine which data sources will be the most appropriate and fruitful, the following requirements should be met; data sources need to:

- **Provide the needed information**.

  For instance, if a pertinent question is whether the program services were accessible to those in greatest need of these services, program staff may <u>not</u> want to conduct a satisfaction survey of clients who have received these services. While such a survey may shed some light on how easy or difficult it may have been for clients to access the services, those surveyed may be the clients who have the least problems with accessibility since they have managed to acquire services. It may be better to identify representatives from the target population and interview them or engage them in a focus group to discuss the particular problems that may exist in accessing services. Or, program staff may want to ask clients who do use services to not only comment on accessibility issues they may have had, but to think about their family and acquaintances who may need these services, but who do not receive them because of access issues. In this sense, the clients who access the program are used as key informants about others who need services, but who do not manage to access them for reasons that these informants can share.

- **Be feasible, given the available human, financial, and other resources.**

  Process evaluation data do not necessarily need to be gathered by evaluation experts. Program staff may be trained to collect and even interpret data in the course of implementing programs. For instance, short questionnaires about what clients appreciated about a service, as well as what they might have liked to receive but did not, could easily be administered by program staff at the end of an individual or group sessions with clients. Further, peers may take turns observing one another in the provision of program services, offering one another feedback on whether services were delivered according to standards and guidelines, and about what could be done differently in the next service-delivery circumstance.

In this case, an outside expert might be used to train peers in carrying out such an assessment process, but not to actually implement the evaluation.

- **Offer confidence in the quality of information gathered.**
  In other words, buy-in in the data-collection methods should be assured. If program staff or stakeholders believe that program feedback was offered only by clients or staff who were "hand-picked" because they were the most satisfied with or vested in the program, then the objectivity of this information may be in question. In this case, feedback should be solicited from randomly selected clients and staff, or, at the very least, carefully chosen respondents who will likely represent a range of perceptions and experiences.

Outcome Evaluation

Because outcome evaluations test whether or not achieved outcomes were due to the program and its interventions or to outside influences, more rigorous study designs are necessary. As such, outcome evaluation requires a comparative element in its design. An outcome evaluation may be guided by one of three types of designs:

- *Experimental designs* involve comparisons of clients randomly assigned to a program (experimental group) with others with similar characteristics who are assigned to a control group. For a number of reasons, including cost, ethical considerations of withholding program services from those who might need them, and difficulty in identifying true control groups, experimental designs may not be possible for outcome evaluations concerning HIV/AIDS-related programs.

- *Quasi-experimental designs* roughly replicate experiments by comparing those individuals who receive program services with those who, through a natural or non-randomly assigned process, do not receive the same services. (They may not receive services at all, or may receive alternative services.) When reasonable comparison groups are available, quasi-experimental designs may be very useful in assisting program staff in understanding program effects on clients.
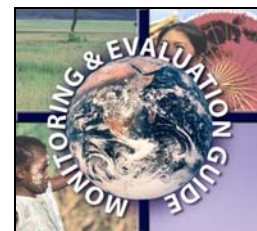
- *Non-experimental designs* may compare clients before and after program participation or over a period of time during participation to learn more about the effects of the program on these individuals. In the absence of a comparison group, non-experimental designs may offer a reasonable way to assess program outcomes. Though these designs do not replicate a scientific experiment, they should not be mistaken as lacking rigor. Well-formulated evaluations using non-experimental designs can complement outcome-monitoring efforts by providing insight that is not possible through outcome monitoring alone.

Since true experiments are less likely to occur in program evaluation, we discuss below several evaluation designs using non-experimental and quasi-experimental approaches. This information is intended to be a primer to assist in beginning to think about program evaluation.

- **Pre-test/Post-test Design:** This approach involves measurement taken of clients prior to a program intervention (pre-test) and again after the intervention (post-test). In this and subsequent examples, O_1 represents a first measurement (or observation) made on a client, O_2 represents a second measurement, and X represents the application of the intervention or service being evaluated. In a pre- post-test design, the sequence is:

$$O_1 \qquad X \qquad O_2$$

This design is most useful in measuring changes in knowledge, skills, attitudes, and self-reported behaviors before and after an intervention; and permits you to compare indicators or measures of the same program participants at two points in time.

- **Retrospective Design:** This design takes a measurement only at one point—after the program intervention. The sequence is:

$$X \qquad O_1 - O_2$$

In this case, O_1 is a retro-measurement—for instance, the client provides a self-report of information about a variable as it existed before he or she received the intervention or service. Immediately after the retro-measurement, in O_2, the client provides a self-report of the variable after receiving the intervention or service. For instance, a client may say, "I didn't understand the behaviors that placed me at risk of HIV infection, but after I received an education session, I understood the sex- and drug-related behaviors that could place me at risk." (Part of the measurement might entail having the respondent list these behaviors to test whether or not he or she now has the correct information.) Retrospective designs are useful when there are time or access constraints that may give you only one chance to gather data from each client. This approach is also beneficial on "highly tested" individuals, whose time you may want to protect. The drawback, of course, is that you are typically dependent on clients' self report for the retro measurement.

- **Time Series:** Time-series depends on a pre-intervention or baseline measurement, then a number of similar measurements after an intervention or service has been delivered. The sequence is:

$$O_1 \qquad X \qquad O_2 \qquad O_3$$

This design allows evaluators to see not only the possible effects of an intervention soon after it is received by a client, but at another time period after the intervention or service is received. This is a desirable approach when follow-up access to clients is not likely to be a problem and when the interest is in understanding of longer-lasting effects of a program, not

only an immediate outcome (e.g., retention of knowledge or skills three months after a program intervention or service was given).

- **Interrupted Time Series:** This is similar to the time-series design, but it depends on multiple measurements taken on the same clients before and after an intervention or service is received. The method uses one group as its own comparison at multiple points in time. The sequence is:

$$O_1 \quad O_2 \quad O_3 \quad O_4 \quad X \quad O_5 \quad O_6 \quad O_7 \quad O_8$$

Compared to the time-series design, interrupted time series offers more stability of the measure over time. That is, you are able to see if the measure fluctuates from measurement to measurement before and after the intervention. If the measure is relatively stable before the intervention, and stable again after the intervention, then you have the ability to make a fair assessment of the effect of the intervention on clients. For instance, if knowledge about HIV risks was minimal in your first measurement and remained that way until the intervention, then knowledge increased in the first measurement after the intervention and remained so at each measurement thereafter, you could state with some certainty that the intervention made a difference in knowledge level.

- **Two-Group, Pre-Test/Post-Test Non-Equivalent Comparison Group:** This design closely replicates an experimental design in that it depends on baseline or pre-intervention (pre-test) and follow-up measurements (post-test) taken from a group receiving program services and a group that does not receive services. The sequence is:

Program group:
$$O_1 \quad X \quad O_2$$
Comparison group:
$$O_1 \qquad\quad O_2$$

Despite the limitation of not having a control group that is completely equal in all characteristics to the program group, it is important to establish reasonable similarity between the two groups in terms of demographics or other factors that are relevant to the group members (e.g., number of children, frequency of unsafe sex). Furthermore, program and comparison group participants should be tested in the exact same way (e.g., using identical measurement instruments) and on the same schedule (e.g., pre- and post-intervention measures are obtained from the comparison group members on the same day or within the same week as the program group).

The effectiveness of the intervention in this design is calculated by the comparison of the difference between the baseline and follow-up measures from the program group, as well as the difference between the baseline and follow-up measures from the comparison group. The primary limitation imposed by this design is that without a true control group, one can never be completely certain that factors other than the intervention produced some of the effects seen (or not seen, as the case may be).

- **Two Group, Interrupted Time Series:** A comparison group could also be used to administer an interrupted time-series design. The sequence is:

Program group:
$$O_1 \quad O_2 \quad O_3 \quad O_4 \quad X \quad O_5 \quad O_6 \quad O_7 \quad O_8$$

Comparison group:
$$O_1 \quad O_2 \quad O_3 \quad O_4 \quad \quad O_5 \quad O_6 \quad O_7 \quad O_8$$

This is the most robust design short of a true experiment in that it combines the stability offered by multiple measurements over time with the ability to compare the group receiving the program intervention or services with a similar group that does not receive these services.

Evaluation Methods

Both process and outcome evaluations may use quantitative or qualitative methods, or a combination of these two general methods. In program evaluation, *quantitative methods* usually involve surveys or other types of questionnaires that gather respondent information that can be aggregated and analyzed. These data often provide the basis for comparing program processes or effects across various types of respondents and/or over various time periods before and after program interventions are delivered. Good instruments ensure validity, reliability, intelligibility, and relatively high response rates. Before conducting surveys—especially those used in outcome evaluations—issues such as adequate sample size, power calculations to assure adequate samples, and handling missing data should be considered.

Qualitative methods such as interviews, focus groups, direct observation, and abstraction of written documents (such as program records) can provide an understanding about social situations and interaction, as well as people's values, perceptions, motivations, and reactions. Interviews typically entail open-ended conversations with program clients and other stakeholders and key informants. A series of consistent questions are usually asked of respondents, but respondents are allowed to answer and elaborate freely through conversations with interviewers. A focus group gathers a small number of individuals (e.g., 5 to 11 individuals per group) to explore ideas, attitudes, experiences, and opinions about a program or service. A focus group is made up of a representation of a targeted demographic group. This method is unique in that it allows for group interaction that provides greater

Quantitative Methods	
Definitions	• Surveys and questionnaires to systematically collect information for a carefully selected sample of individuals or households.
Uses	• Comparing different groups at a point in time. • Comparing changes over time in the same group. • Describing conditions or context. • Providing data for evaluating achievement of outcomes/effects.
Advantages	• Findings can be generalized to the larger population represented by respondents if certain conditions in the evaluation design are met. • Provides basis for quantitative estimates of size and distribution of outcomes and impact.
Disadvantages	• Data collection, analysis, and processing may take some time; lack of immediate data and findings to use. • May be expensive for large surveys. • Many kinds of information are difficult to obtain through surveys.
Costs/Time	• Medium to high cost, depending on sample size and complexity of survey. • Data collection, processing, and analysis can take anywhere between 2 months and 2 years, depending on size of sample.
Skills	• Good technical and analytic expertise for determining samples, questionnaire design, and data analysis and processing.
Types/Examples	• Mail, telephone, Web-based, polls. • Knowledge, attitude, behavior. • Satisfaction. • Household information.

insight into why certain opinions or ideas are held. A skilled focus group facilitator assures that the group stays on target in terms of providing the kind of information needed by program evaluators, and uses the group process itself to gain depth of insight in a way that is more beneficial than attempting to obtain this information from individual respondents. Discussion groups are less structured than focus groups and serve to gather information about programs from larger numbers of people or individuals representing a range of program stakeholders.

Case studies and site visits often entail a combination of qualitative methods. (See Table 1 below for example of case studies.) These approaches entail on-site visits to programs or organizations to conduct interviews with individuals or groups, and to gather other data *in situ* (i.e., in the context in which programs are implemented). Five keys to conducting successful site visits are: selecting the right site, allaying the site's fears about being scrutinized and getting its buy-in, minimizing disruption of their activities, incorporating benefits to the site for their participation, and ensuring the utility of collected information for the site itself. Site selection can be straightforward once appropriate selection criteria are defined. For example, some evaluation projects require exemplary or illustrative sites, while others need a mix that provides representation of particular characteristics. Site visits often involve personal interviews or focus groups, program document review, observations, and when appropriate, workflow analyses.

Qualitative Methods	
Definitions	• Interviews, focus groups, direct observation, and abstraction of written documents (such as program records) can provide an understanding about social situations and interaction, as well as people's values, perceptions, motivations, and reactions.
Uses	• Providing information for management decision-making at project/program level. • Providing qualitative understanding about complex social situations and interaction, as well as people's values, perceptions, motivations, and reactions. • Providing context and interpretation for quantitative data collected with other methods.
Advantages	• Usually low cost. • Usually relatively quick. • Provides opportunity to explore new ideas.
Disadvantages	• Findings seldom able to be generalized. • May be less valid, reliable, and credible than survey methods, depending on methods used.
Costs/Time	• Low to medium cost (ethnographies can be costly, however). • Data collection, processing, and analysis may be time consuming, depending on methods used.
Skills	• Interviewing. • Group facilitation. • Observation. • Note taking. • Basic statistical skills for limited data that can be quantified.
Types/Examples	• Stakeholder/key informant interviews. • Focus groups. • Discussion groups/community group interviews. • Participant observation. • Site visits. • Ethnographies. • Mini-survey.

Table 1. Example of Using Case Studies

I. Background and Methods

Background:

The Zimbabwe CDC/GAP (Zim-CDC) country program began in December 2000. A year later, as part of its overall M&E strategy, the CDC/GAP M&E team visited the program to obtain an in-depth description of GAP activities in Zimbabwe. A case-study methodology was used to collect information from country partners working with the Zim-CDC program, as well as other stakeholders. Interviews focused on information about ZIM-CDC's role in:

(1) Fostering collaboration among public health agencies and other partners.
(2) Enhancing the quality, availability, and use of surveillance and other data, including increasing capacity for M&E.

Case Study Methods:

- The case study was jointly conducted by a member of the CDC/GAP M&E Team and an external advisor from UNAIDS who was familiar with the Zimbabwe context. The CDC/GAP investigator made two field trips to Zimbabwe (September 29 to October 12, 2001; November 9 to November 24, 2001)
- The first field trip was to become familiar with the CDC program in Zimbabwe through document review, participation in meetings and activities as part of the daily activities of the Zim-CDC office, and interviews with Zim-CDC staff.
- The second field trip aimed to obtain views from stakeholders and partners regarding Zim-CDC's approach and program. The CDC/GAP investigator and UNAIDS consultant conducted interviews. Interviewees were selected based on their knowledge of and/or involvement with the Zim-CDC program. To minimize bias, at least one representative from each stakeholder group, collaborating partner organization, and major donor agency was selected. Interviewees were asked to advise on any other key individuals who could be included in the study. Standard qualitative research protocol was followed, including:
  - ✓ Interviews were tape-recorded and written notes were taken during interviews.
  - ✓ Written informed consent was gained from interviewees, including guarantee of confidentiality and that interview tapes and notes would be kept secure.
  - ✓ Before finalization of the case-study report, interviewees who requested to comment on the draft report were given the opportunity to do so, and revisions were made accordingly.
- Interviews were conducted with 31 stakeholders and partners representing 17 different agencies/institutions. Types of agencies included:
  - ✓ Government of Zimbabwe authorities. (2 agencies, 10 interviews)
  - ✓ Local agencies/institutions. (7 agencies/institutions, 9 interviews)
  - ✓ Multi-lateral and bi-lateral agencies. (6 agencies, 9 interviews)
  - ✓ USG agencies. (2 agencies, 3 interviews)
- Interview questions focused on:
  - ✓ Initial contacts with CDC. (When did they place? What was the purpose? What were the ensuing expectations)
  - ✓ Continued communication with CDC. (What is the regularity and nature of these interactions?)
  - ✓ Appropriateness of the CDC program. (What is CDC's proper role? Does the CDC program fit with the national response to HIV/AIDS? Is CDC's program complementary to other ongoing programs?)
  - ✓ Satisfaction with CDC's progress to date.
  - ✓ Any concerns/challenges. (Lessons learned from working with Zim-CDC.)

Table 1. Example of Using Case Studies (cont.)

II. Findings

The following summarizes findings from the Zim-CDC case study. Stakeholders'/partners' views, and not those of CDC staff, are represented in these findings:

Achievements:

(1) Fostering collaboration among public heath agencies and other partners.

- The initial assessment visits made by CDC and the and extensive consultations with local stakeholders to discuss CDC's role and develop the ZIM-CDC program were highly valued.
- Zim-CDC has achieved high credibility in its assistance with the public health aspects of the national response to HIV/AIDS.
- The Zim-CDC program is responding to needs identified in the national strategic plan. Zim-CDC has a strong commitment to working within existing government systems, rather than parallel systems.
- The Zim-CDC program is perceived to be appropriate and complementary to existing programs.
- Zim-CDC staff's enthusiasm is an important catalyst in the national response to HIV/AIDS.
- Zim-CDC is perceived to play an important role in strengthening coordination between different players on behalf of the government of Zimbabwe.
- There is a high level of satisfaction with Zim-CDC's responsiveness in terms of timely response, program transparency, openness towards stakeholder/partner opinions, and flexible procedures.

(2) Enhancing the quality, availability, and use of surveillance and other data, including the strengthening of M&E capacity.

- Surveillance and M&E are priorities, and capacity-building for these activities is urgently needed.
- Zim-CDC's assistance with surveillance and M&E is timely.
- Zim-CDC is technically competent and results-oriented.
- Zim-CDC draws on local expertise in program planning and implementation.
- Zim-CDC's support is highly valued in strengthening information technology and systems.

Challenges:

- The diversity of the Zim-CDC program must be balanced with sufficient coverage and depth to maximize its impact.
- Although the response to the HIV/AIDS crisis is urgent, enough time needs to be taken to ensure that all stakeholders are "on board." The challenge is to balance results-oriented and process-oriented approaches.
- Although a need clearly exists to build capacity in M&E at different levels, responsibilities must be clearly delineated under the leadership of the National AIDS Council.
- Zim-CDC should encourage the NAC to use surveillance findings more aggressively to gain the momentum of support and involvement of civil society on HIV/AIDS issues.
- The expertise of nurses and midwives should be drawn into the Zim-CDC program.
- Opportunities for exchange of information and lessons learned should be increased within the Zim-CDC program and between countries.
- Although Zim-CDC's support in informatics and building information systems is crucial, its approach may be too ambitious, and it should remain flexible to allow for adjustments.
- Accessing CDC funds by means of grants is a lengthy procedure and must be balanced with adequate direct funding.
- A broad consultation among donor agencies is needed to formulate a consistent policy on supporting salaries for local positions.

Lessons Learned:

- Conducting a needs assessment and consulting with key stakeholders in advance of program initiation helps to ensure that the program responds to actual country needs and defines an appropriate role for an agency beginning work in a country. Time is well spent in researching the current activities and learning who is doing what before establishing a new program.
- Although many agencies and organizations may be already working in a country, CDC/GAP can bring a new and re-vitalizing perspective to combating HIV/AIDS.
- A resident CDC/GAP program office is essential for establishing and maintaining good working relationships and in getting things done.
- CDC/GAP is most valuable in contributing to the health aspects of the HIV/AIDS response based on its public health vision and expertise.
- Striking a balance is crucial between supporting a broad program and obtaining sufficient depth and coverage to achieve maximum impact.
- As essential component of CDC/GAP approach is its collaboration with government agencies to strengthen the much-needed leadership and coordination of the national response to HIV/AIDS.
- Setting clear priorities and leveraging long-term donor support are key to capacity-building.
- Mechanisms for efficient country-to-country sharing need to be developed.

Further Tools and Reading

Chapel Hill, NC, and UNAIDS, Geneva, Switzerland. Available online at:

http://www.cpc.unc.edu/measure/publications/workingpapers/wp9917.pdf

Campbell, D.T., and Stanley, J.C. (1966). Experimental and Quasi-Experimental Designs for Research. Houton Mifflin.
A classic resource that covers experimental and quasi-experimental research designs, as well as those that are not adequate.

Mantell, J., DiVittis, A.T., and Auerbach, M.I. (1997). Evaluating HIV Prevention Interventions. New York, Plenum Press.
See Chapters ("Selecting an Evaluation Study Design"), Chapter 6 ("Quantitative Measures in Evaluation"), and Chapter 7 ("Using Qualitative Methods").

Peersman, G., and Sikipa, G. (2002). *Strengthening the National Response to HIV/AIDS. The Role of the CDC Global AIDS Program in Zimbabwe: A Case Study.* U.S. Centers for Disease Control and Prevention, Atlanta, GA.

Rugg, D., and Mills, S. (2000). Development of an integrated monitoring and evaluation plan. In Rehle, T., Saidel, T., Mills, S., Magnani, R. (Eds*.), Evaluating Programs for HIV/AIDS Prevention and Care in Developing Countries.* Family Health International, Arlington, VA.

Rugg, D., Buehler, J., Renaud, M., et al. (1999). Evaluating HIV Prevention: A Framework for National, State and Local Levels. *American Journal of Evaluation* 20, 35-56.

United Nations Population Fund. (2001). Tool No. 5, Planning and Managing an Evaluation, Part III: The Data Collection Process. *Monitoring and Evaluation Toolkit for Programme Managers.* Office of Oversight and Evaluation:

http://www.unfpa.org/monitoring/toolkit/5evaluation.pdf
Available in English, French, and Spanish.

Adapted in part from *Monitoring & Evaluation: Some Tools, Methods, & Approaches*, June 2002, Washington, D.C.: The World Bank.

9. DATA MANAGEMENT AND ANALYSIS

It is likely that program staff will need appropriate staff or individualized TA for some data management and analysis issues. For instance, as programs grow, they may find it beneficial to have computer software programs and data management systems that meet the needs of particular program staff and clients. But these programs and systems should have some ability to relate to larger data management information systems that have been developed by GAP and other global partners to track and analyze national- and global-level indicators. Not all programs may need these; and for programs that do, program staff are not expected to create these systems in isolation. These staff will have the benefit of resources from GAP HQ and other country field staff who have already developed systems that work for them.

However, many individuals and programs have more inherent capability than they initially believe to conduct relatively straightforward to moderately complex data management and analysis activities. Below we provide some basic information for getting started with data management and analysis. Again, this is only a beginning, and most of the information, skills, and technology needed will be best met through appropriate staff or TA provision.

Data management systems and methods of analysis should be considered <u>before</u> data collection begins. Program staff should be prepared to store and process data as this information comes in; otherwise a glut of paper-generated data will clutter program offices before analysis and management plans are developed. To begin this plan, evaluation questions may be reviewed to determine the types of data and analyses that will be required. Each question, reviewed individually, should help program staff to conceptualize the analysis and data storage and processing needs.

Data Analysis

As an overview regarding data analysis, quantitative data require quantitative data analysis methods; qualitative data require

qualitative data analysis approaches; and multiple methods require both of the appropriate types of analysis relative to the type of data, but may also necessitate extra steps in linking one type of data to another.

Factors influencing the choice of quantitative analytic approaches include the research questions, sample size, data structure, software availability, and statistical assumptions. Questions may be answered with approaches ranging from simple descriptive statistical analysis, t-test, and cross-tabulation procedures, to analysis or multivariate analysis of variance or co-variance. For most program M&E activities, simple statistics and, perhaps, some cross tabulations are likely all that will be required. Aside from the very basic discussion below about developing simple statistics, your needs about data analysis will be addressed through individualized TA. Rarely will projects need to employ analysis of variance and co-variance. For outcome evaluation designs that may benefit from such analysis, TA may be requested as well. Other, more complex statistical methods (e.g., mediational analysis and complex theoretical modeling) are typically not part of program evaluation; therefore, they will not be discussed here.

Qualitative analysis should be no less rigorous and may have as many considerations as the analysis of quantitative data. Such techniques and methods may include: following inductive analysis with a search for rival or competing themes and explanations; exploring variation and searching for negative cases; triangulating methods, sources, theories, or perspectives to test for consistency and reduce systematic bias; and, keeping methods and data in context so as to avoid overgeneralization of the findings. Just as new statistical techniques and software programs have been developed for the analysis of quantitative data, new approaches and software programs have also emerged for qualitative data. For example, Nudist and Atlas-ti are two useful programs currently on the market. The CDC-developed Analysis Software for Word-Based Products (Answr) is also available at: http://www.cdc.gov/hiv/software/answr.htm.

For multiple-method studies, appropriate analytic methods should be used that match each type of data collected. Data emerging from multiple methods also may be linked through analytic approaches to arrive at a richer and deeper understanding of the phenomena under study. A good "primer" on methods for linking qualitative and quantitative data is: Fielding N.G., & Fielding, J.L. *Linking Data*, Newbury Park, CA, Sage, 1989.

Data Management

It is important that data be stored in a safe and protected place that will ensure that important information will not be destroyed in rare instances such as fires and floods, and that client information is kept confidential. To guarantee client confidentiality, client names and other identifying information should be kept separate from sensitive information, such as HIV status and history of STI. Usually, a coding system that assigns a numerical or alphabetical code to each client is sufficient. Client names and assigned codes should be placed in a hardcopy or electronic file separate from the actual M&E data to assure confidentiality.

A variety of electronic data management systems exist to facilitate data storage and processing. Though field offices will want to use electronic systems, it is possible that various remotely located field projects that are part of your programs will not have the technological capacity to provide electronic data. In these cases, field office staff may need to find a way to have hardcopy field data entered into the electronic systems at the field office.

If computer software to manage and process data has not been obtained, program staff may want to hire appropriate staff or obtain individualized TA to determine your system needs. Also, field office staff may want to seek advice from peers at other field offices that have established data management systems. Below are a number of the more accessible software packages available:

Epi Info

Software available **FREE** at: http://www.cdc.gov/epiinfo

Using Epi Info 2002: *A Step-by-Step Guide* by C. Escofrey, M. Alperin, & K. Miner

- Publisher: Toucan Ed; Book and CD-ROM (December 2002)
- To order**:** http://www.toucaned.com

Microsoft Access 2002

Microsoft Access Version 2002 Step by Step by Online Training Solutions, Inc.

- Publisher: Microsoft Press; Book and CD-ROM edition (June 6, 2001)
- ISBN: 0735612994

Microsoft Access 2002 for Dummies by John Kaufeld (Author)

- Publisher: John Wiley & Sons; 1st edition (January 15, 2001)
- ISBN: 0764508180

Microsoft Excel 2002

Microsoft Excel Version 2002 Step by Step by Curtis Frye

- Publisher: Microsoft Press; Book and CD-ROM edition (June 6, 2001)
- ISBN: 073561296X

Excel 2002 for Dummies by Greg Harvey (Author)

- Publisher: John Wiley & Sons; (June 2001)
- ISBN: 0764508229

FileMaker Pro

Learn FileMaker Pro 6 by Jonathan Stars, Nonie Bernard

- Publisher: Republic of Texas Press; Book and CD-ROM edition (December 2002)
- ISBN: 1556229747

Further Tools and Reading

In addition to the above resources, the following are suggested:

Barry, C. (1988). Choosing Qualitative Data Analysis Software:
Atlas/ti and Nudist Compared. *Sociological Research Online* 3,3:
http://www.socresonline.org.uk/socresonline/3/3/4.html

NIDA. (1991). *Handbook for Evaluating Training in HIV/AIDS and Illicit
Drug Use: A Manual for State Drug Use Authorities*. Community and
Professional Education Branch, National Institute on Drug Abuse
(also see Program Evaluation for STD/HIV Behavioral
Interventions, California STD/HIV Prevention Training Center).

10. REPORTING AND USING DATA

Reporting and Using Data

Besides reporting of data through Annual Reports to CDC/GAP HQ, it is important to disseminate key findings from M&E among the stakeholders and users who assisted in determining the evaluation

<table>
<tr>
<td>

Possible uses of findings by the program being evaluated:

- Informing strategic program planning for the future.

- Informing capacity-building plans and activities.

- Informing program staffing plans and decisions.

- Informing funding efforts and decisions.

- Affecting development of and changes to organizational policy.

Possible uses of findings by outside programs and agencies:

- Providing valuable information about lessons learned for agencies planning to implement similar programs.

- Affecting government policy and procedures.

</td>
<td>

questions. Some of these findings may overlap with required data reporting, but they will also include other program-improvement information gathered apart from required reporting. Findings may have implications for continuation and revisions to the program that was evaluated, as well as for agencies interacting with this program and other programs hoping to implement similar processes and approaches. Though there are many possible uses for evaluation findings, the textbox to the left illustrates some of the most typical uses:

Table 2 provides examples of select findings from the Zimbabwe/CDC Case Study introduced in Chapter 8.

Before deciding to share findings, it is important to determine whether the evaluation actually produced credible information based on sound methods and recommendations that are pertinent to key stakeholders. To assess whether evaluation findings are ready for dissemination, the following checklist of important evaluation standards should be satisfied:

- The evaluators were credible, appropriate for the type of evaluation conducted, and did not exhibit a conflict of interest (See Chapter 11 on choosing evaluators).
- Information gathered was valid, or measured what was intended to be measured, and reliable, or could produce the same results repeatedly.

</td>
</tr>
</table>

- The evaluation process was transparent. Rationale and procedures used were clearly explained and limitations faced by the evaluation were described.
- Standard research ethics were followed, including protection of human subjects.
- The evaluation produced a clear and fair assessment of the question under study. Weaknesses as well as strengths of the program were discussed and the context in which the program took place was explained.
- Information sources were appropriate and biases were explained.
- Systematic data collection, storage and cleaning, and analysis was followed, and systematic review of each stage of evaluation was carried out.
- Analytic techniques were used appropriate to the evaluation question(s) and methodological standards (see Chapter 8 on Evaluation Methods).
- Conclusions were justified.
- Recommendations were appropriate to findings and pertinent to stakeholder needs.

A full list and explanation of evaluation standards may be found in Part VI of the United Nations Population Funds, *Monitoring and Evaluation Toolkit for Programme Managers* (full reference below).

A number of avenues exist for disseminating findings. To begin thinking about the best ways of sharing your information, consider existing formal and informal networks within the communities in which programs take place. These include community and peer meetings scheduled specifically to review findings, as well larger meetings within which a special time is set aside to discuss findings. Local newsletters and other types of regular communication through which programs already share information may be natural places to post key findings. Also, professional conferences encourage presentation of M&E results as discussion papers or posters. These conferences often make abstracts available to an even wider audience. Another avenue to disseminate findings is through professional or lay

journals if, through M&E data, lessons learned would benefit a larger audience. Key findings may be shared electronically through appropriate Web pages, as well as through Web-based information clearinghouses.

Written reports should be constructed in a way that best suits needs of the intended audience and existing guidelines established by the media source. As a general rule, most evaluation reports include:

- An executive summary.
- A section on background or context that describes the program and the problem it addresses.
- A description of data collection and analysis.
- Summary of the M&E results.
- Conclusions and recommendations.

Table 2: Example of Using Evaluation Findings: Select Findings from ZIM/CDC Case Study

The following are select findings from a 2001 case study of the Zimbabwe CDC/GAP (Zim-CDC) country program begun in December 2000. In interviews, information was collected from country partners working with the Zim-CDC program, as well as other stakeholders. Interviews and resulting findings focused on (1) fostering collaboration among public health agencies and other partners, and, (2) enhancing the quality, availability, and use of surveillance and other data, including increasing capacity for M&E. (See Table 1, Chapter 8, for more information about the background and methods used in this case study.)

Using this table: Findings from this case study appear in the left-hand column of the table. For each finding, one or more implications are marked with an X in the remaining columns. For instance, for the first finding shows that Zim-CDC's has a strong commitment to working with existing government systems. An implication of this finding is that working with government systems rather than constructing a parallel system is a fundamental strategy that informs Zim-CDC's strategic program planning.

Implications for Continuation and Change

SELECT FINDINGS	Strategic Program Planning	Capacity-building	Staffing plans	Funding	Organizational Policy	Planning Similar Programs	Government Policy & Procedures
The Zim-CDC program is responding to needs identified in the national strategic plan. Zim-CDC has a strong commitment to working within existing government systems, rather than parallel systems.	X						
Zim-CDC is perceived to play an important role in strengthening coordination between different players on behalf of the government of Zimbabwe.		X					
Surveillance and M&E are priorities, and capacity-building for these activities is urgently needed, and Zim-CDC's assistance with these areas is timely.		X	X				
Zim-CDC's support is highly valued in strengthening information technology and systems.	X						
The diversity of the Zim-CDC program must be balanced with sufficient coverage and depth to maximize its impact.	X		X			X	
Zim-CDC should encourage the NAC to use surveillance findings more aggressively to gain the momentum of support and involvement of civil society on HIV/AIDS issues.							X
The expertise of nurses and midwives should be drawn into the Zim-CDC program.			X				
Accessing CDC funds by means of grants is a lengthy procedure and must be balanced with adequate direct funding.				X			
A broad consultation among donor agencies is needed to formulate a consistent policy on supporting salaries for local positions.			X		X		X
CDC/GAP is most valuable in contributing to the health aspects of the HIV/AIDS response based on its public health vision and expertise.	X		X			X	
Setting clear priorities and leveraging long-term donor support are key to capacity-building.		X		X		X	

Further Tools and Reading

Peersman, G., and Sikipa, G. (2002). *Strengthening the National Response to HIV/AIDS. The Role of the CDC Global AIDS Program in Zimbabwe: A Case Study.* U.S. Centers for Disease Control and Prevention, Atlanta, GA.

United Nations Population Fund. (2001). Tool No. 5, Planning and Managing an Evaluation, Part V: Communicating and Using Evaluation Results; Part VI: Evaluation Standards. *Monitoring and Evaluation Toolkit for Programme Managers.* Office of Oversight and Evaluation:

http://www.unfpa.org/monitoring/toolkit/5evaluation.pdf

Available in English, French, and Spanish.

United Nations Population Fund. (2003). *Evaluation Reports and Findings.* Office of Oversight and evaluation:

http://www.unfpa.org/monitoring/reports.htm

Sample evaluation reports, including HIV/AIDS-related evaluation findings and recommendations.

11. EVALUABILITY ASSESSMENT: PROGRAM READINESS FOR M&E

Although GAP program staff are required to monitor their programs and are encouraged to conduct program evaluation, M&E activities are not necessarily easy or inexpensive to implement. Before undertaking evaluation activities, it is wise for staff to conduct an evaluability assessment before beginning process or outcome evaluations of programs. An evaluability assessment provides useful information on a program's readiness to be monitored and/or evaluated. If a program is not ready for evaluation, limited funds, time, and other resources might be wasted and program staff might be discouraged from conducting evaluation. This chapter will detail the steps to assessing a program's evaluability.

Evaluability Assessment

To determine program readiness for monitoring and/or evaluation, a series of specific questions must be answered. When the answer to each question is "yes," then you are ready to evaluate. A "no" answer means that some action should be taken to better prepare for evaluation. For each "no" response, brainstorm actions should be taken to guide future efforts to prepare for evaluation. The following questions gauge a general readiness to evaluate:

- Is there a willingness to monitor/evaluate?
- Have the intended users of the evaluation been identified?
- Is there a logic model describing planned implementation and outcomes?
- Have evaluation questions been identified?
- Is there a desire to use the evaluation findings?
- Have data needs been determined?
- Are the data needed available or feasible to collect?
- Have evaluation resources been secured?

- **Sustainability:**
  The political and financial will exist to sustain the program while the evaluation is conducted.
- **Fidelity:**
  Actual program implementation matches intended implementation plan, determined via program monitoring or process evaluation.
- **Stability:**
  Program is not likely to change during the life of the program or during the program/intervention period being evaluated.
- **Reach:**
  Program reaches a sufficiently large number of clients (sample size) to apply statistical tests necessary for data analysis.
- **Dosage:**
  Clients have sufficient exposure to the program to result in the intended outcomes.

- Have data collection, management, and analysis procedures been developed?
- Is there a strategy to disseminate and use the evaluation findings?

If your program is preparing to conduct <u>outcome</u> monitoring and evaluation, there are some additional questions to consider that do not apply to input/output monitoring and process evaluation.

- Is the program sustainable?
- Is the program implemented with fidelity to its plan (including links between goals, objectives, and activities)?
- Is the program stable over time?
- Does the program reach a sufficient number of people?
- Is the program delivered with sufficient dosage?

Further information on the variables listed in these questions may be found in the textbox on the previous page.

Building M&E Capacity of Existing Staff

Technical Assistance (TA) Resources

- National and local consultants
- University TA Providers (UTAP)
- GAP M&E Team
- University faculty
- Graduate students
- Volunteers (e.g., evaluation association members

Once it is decided that a program is ready for evaluation, program staff must assess if they have the necessary capacity to conduct effective monitoring and evaluation and use M&E findings for program management and improvement. Important questions to ask about staff capacity include:

- Are there enough staff to conduct M&E?
- Do staff have experience in M&E?
- What do we want to know about our programs?
- How will we collect M&E information?
- How will we store and analyze M&E information?
- How will we use M&E findings?

Based on the answers to these questions, programs might decide to conduct M&E activities using existing staff, use external evaluators for discrete M&E activities, or hire new staff dedicated to M&E activities. TA providers are available to support and collaborate with programs conducting M&E (see text box to the left).

Working with TA providers requires establishing and maintaining effective working relationships with the individuals or organizations providing the TA. When working with evaluation consultants, it is important to select a TA provider who knows the topic well, is culturally competent, and can communicate clearly with different stakeholders, such as program managers, funders, and community members. From the beginning, program staff should clarify the roles and responsibilities of the TA provider, establish a work plan and timeline, and schedule regular meetings with the TA provider to monitor progress. Finally, working with a TA provider creates a unique opportunity to build internal M&E capacity for future M&E activities.

Choosing and Using External Evaluators

Sometimes it makes more sense to contract with external evaluators to obtain objective, well-planned, and well-conducted program evaluations. It is critical that program staff take time to select the most suitable evaluator who possesses the necessary skills for M&E, understands the program to be evaluated, and designs evaluation approaches providing useful and appropriate information. The following guidelines present important considerations for selecting and working with an appropriate external evaluator. (Note: These guidelines were adapted from the *Handbook for Evaluating HIV Education*, Division of Adolescent and School Health, Centers for Disease Control and Prevention.)

- **Use an evaluation liaison or committee.**

  An evaluation liaison or committee of program staff is responsible for the evaluation and coordinates with the external evaluator. The liaison should be involved in the evaluation process to increase understanding of the evaluation design and methods, and to ensure usability of evaluation results.

- **Define the evaluation.**

  The evaluation liaison or committee must focus the evaluation with specific evaluation questions to be answered and tasks to be completed by the external evaluator. Examples of tasks for which an external evaluator would be responsible include:

  – Development of the evaluation plan and instruments.
  – Selection of sampling procedures.
  – Drawing of the evaluation sample.
  – Evaluation data collection and analysis.
  – Production of the evaluation report.
  – Presentation of evaluation results.

- **Solicit candidates.**

  Once the evaluation has been defined, the evaluation liaison should compose a job description including the identified tasks, other requirements for the position, a description of the program to be evaluated, and an estimate of the funds available for the evaluation. The job description also should include instructions of how candidates can apply for the position, and what information and documentation must be provided to the liaison. Additionally, the evaluation liaison or committee should review at least one evaluation report written by the applicant and at least two references who received evaluation services from the applicant.

  The job description can be distributed to local colleges and universities, professional organizations such as evaluation associations, non-governmental organizations and aid agencies, and existing networks of contacts to identify program evaluators in the community. The job description can also be advertised in national and global public health journals, associations, and job Websites. If external evaluators are selected from outside the local area, ensure that they are able to participate in regular meetings regarding the evaluation or have reliable communication systems.

- **Interview and select the evaluator.**

  The evaluation liaison or committee should narrow the pool of applicants to five or fewer candidates for formal interviews. The following questions would be helpful to explore relevant issues during the interviews:

  − *Does the candidate understand the difference between research and evaluation?* Sometimes evaluators are more interested in conducting research outside the scope of the evaluation, often limiting the utility of the evaluation results. To confirm program evaluation skills and support the use of evaluation results, ask the candidate to describe the difference between research and evaluation.

  − *Does the candidate understand the program to be evaluated?*

  − *What would the candidate's general approach to the evaluation be?* Ask the candidate to describe the general evaluation approach he or she would use. Note the focus of the evaluation questions and issues and how well the candidate presents technical information on data collection and analysis.

  − *Does the candidate believe the evaluation can be conducted with the available funds?* Ask all candidates who participate in a formal interview to produce a detailed budget for their evaluation approach. Ensure that the evaluation plan is possible given available funds.

  − *What is the candidate's reaction to supervision by the evaluation liaison or committee?*

  − *What is the candidate's prior evaluation experience?* Depending on the nature of the evaluation, specific prior experience may include but not be limited to: protocol development; survey design, implementation, and analysis; use of qualitative data collection methods; and knowledge of or ability to incorporate CDC human subjects guidelines and review process. Experience with

field implementation of public health programs with expertise in at least one program area is also desirable.

- *How useful are the candidate's previous evaluation reports?* Assess prior evaluation reports for clarity, organization, readability, and usefulness for decision makers.

- *Does the candidate have good references?* Contact references for objective views of the candidate. Determine if the evaluation approach the evaluator used addressed their needs and desires, if the evaluation was conducted in a timely fashion and within budget, if the report was useful, and if they would use the evaluator again.

- *Will the candidate's existing commitments interfere with the planned evaluation?* If a candidate has several other professional commitments, ask how he or she will conduct the evaluation in addition to the other commitments. If other staff will assist with the evaluation, determine who will be responsible for which tasks and if the other staff are capable of performing the assigned tasks.

- *What is your general reaction to the candidate?*

- *What is your overall rating of the candidate?*

- **Write and negotiate the contract.**

  The contract should state the evaluator's responsibilities, a detailed decision-making process, and the authority of the evaluation liaison or committee. The contract also should include deliverables and a timetable, who owns the data collected during the evaluation, and who has the rights to publish the results of the evaluation study. Additionally, the contract should outline how the evaluator will bill for services rendered and a schedule of payment. Generally, a percentage of the evaluator's fee should be withheld until the evaluation liaison has accepted the final report. Finally, the contract should state the responsibilities of the evaluation liaison to

provide the evaluator with timely guidance, to review and approve instruments and documents, and to assist the evaluator in solving problems if they should arise.

- **Interact closely with the evaluator.**

  At the first meeting with the evaluator, and at regular, subsequent meetings, the evaluation liaison should communicate expectations for the evaluation, review materials for the evaluation, and monitor progress of the evaluation.

- **Prepare the final report and release of results.**

  Towards the end of the evaluation, a report format should be decided upon, and the evaluator and liaison should agree to a plan for release of results. As drafts of the report are prepared, the evaluator should submit them to the liaison for review and comment. Significant changes to the findings, recommendations, or overall focus of the report should be discussed between the evaluator and the liaison.

Hiring M&E Staff

If a program has the resources to support an internal evaluator and the organizational drive to conduct ongoing M&E activities, then a dedicated M&E staff person might be hired. An internal M&E specialist brings M&E expertise to the program, providing guidance to other staff and resources for well-designed and useful M&E activities. The process for identifying and hiring internal M&E candidates with appropriate experience mirrors the above steps for using external evaluators. Prior evaluation and public health program experience is necessary, and the specific requirements listed above as potential requirements for external evaluators should all be included as requirements for internal M&E staff. Again, these include but are not necessarily limited to: protocol development; survey design, implementation, and analysis; use of qualitative data collection methods; and knowledge of CDC human subjects guidelines and review process. Experience with field implementation of public health programs with expertise in at least one program area is also desirable.

> Examples of M&E staff job descriptions and scope of work appear in Appendix J.

Resources for Evaluators

American Evaluation Association webpage:

> http://www.eval

Africa Evaluation Association webpage:

> http://www.geocities.com/afreval

European Evaluation Society webpage:

> http://europeanevaluation.org

Evaltalk:

> http://www.eval.org/ListsLinks/ElectronicLists/evaltalk.html

French Evaluation Society webpage (in French):

> http://www.sfe.asso.fr

Malaysian Evaluation Society webpage:

> http://www.angelfire.com/ab/mes/index.html

Preval for Latin America & the Caribbean webpage (in Spanish):

> http://preval.org

WWW Virtual Library: Evaluation webpage (many international organizations): http://www.policy-evaluation.org

PART III: PLANNING AND REPORTING REQUIREMENTS

12. GAP M&E FRAMEWORK AND EXPECTATIONS

Overall M&E Strategy

GAP M&E goals have been identified as:

Goal 1: Determine the progress and effectiveness of CDC/GAP programs and assistance activities.

Goal 2: Strengthen the capacity of National HIV/AIDS Programs to conduct monitoring and evaluation.

To achieve these goals, nine critical elements form the GAP M&E Strategy as depicted in the figure below.

Figure 15.   Critical Elements of GAP M&E Strategy
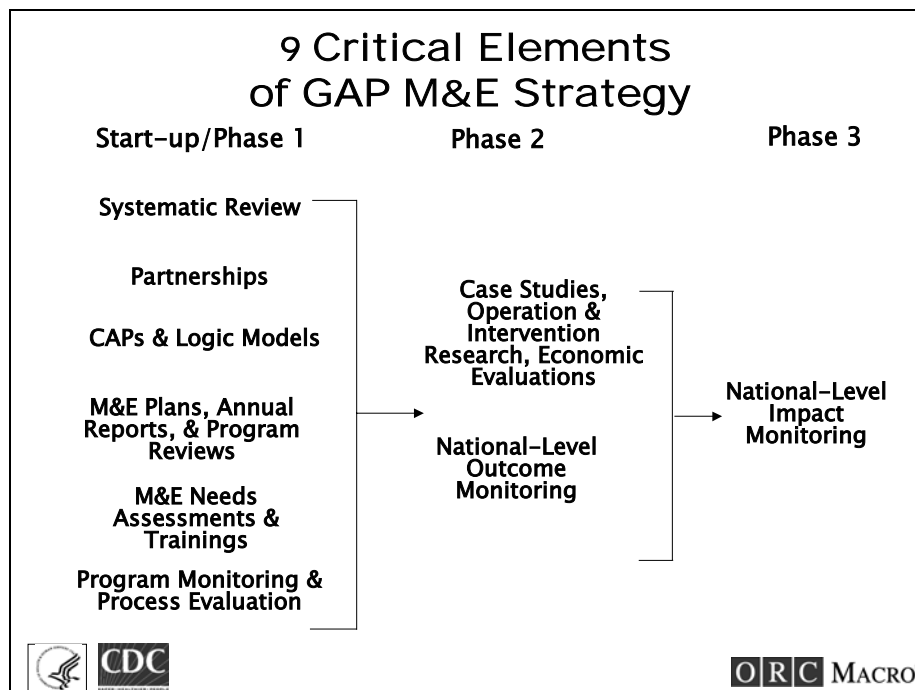
As the model of the strategy shows, the first steps in this three-phased strategy involve planning steps, including the first critical element, the systematic review of existing HIV/AIDS behavioral interventions that are targeted to various populations at risk of or infected with HIV. The goal of this review is to identify successful evidence-based, population-specific interventions that may be replicated in GAP countries. Of course, the generation of information about evidence-based practices requires partners and stakeholders who design

programs, interventions, and related strategies based on this knowledge. Therefore, a necessary next step is to share information with GAP partners and stakeholders through publications, presentations, as well as face-to-face and other forms of communication. Evidence-based information is particularly useful for GAP field offices in developing the third critical element of the strategy, the Country Assistance Plan (CAP) and related program logic models that assist in establishing the logical steps and program needs (e.g., staffing, materials, referral partners). The CAP, which is the document that allows GAP country programs to communicate their respective countries' HIV/AIDS context and plans for contributing to the national response, requires the establishment of program objectives related to program plans and logic models.

The forth critical element of the M&E strategy is development by field offices of M&E Plans and systems for conducting M&E activities, using M&E data for annual reporting, and conducting GAP Country Program Reviews. Field offices are supported by the GAP HQ M&E staff in identifying appropriate M&E approaches, as well as implementing M&E activities. This support begins with the assessment of needs for planning and implementing M&E activities. Identified needs are met through both formal face-to-face training in M&E and on-the-ground TA (fifth critical element in Figure 15). TA that the HQ M&E staff are able to provide is supplemented by a number of university, private, and other TA providers to meet various M&E needs.

Once field offices have identified specific M&E activities in an M&E plan and have been trained in basic M&E knowledge, field office staff are expected to track inputs (i.e., resources put into a program) and outputs (i.e., results of program activities) over the course of program implementation (sixth critical element). Some may also want to assess the quality of the program.

The seventh critical element involves identifying and conducting special evaluation studies, such as:

- Case studies designed for gaining knowledge to fill very specific information gaps about the effectiveness of interventions on particular populations.

- Outcome evaluations or intervention outcome studies that are carefully designed to systematically collect data about whether or not outcomes achieved by a program were due to the program interventions or other outside variables.

- Economic evaluations that assess the cost effectiveness and cost benefits of programs.

The eighth critical element of the M&E Strategy involves collaboration with host governments and other GAP partners in tracking of national-level outcomes. To conduct such monitoring, GAP and its partners have identified a number of indicators that, when measured, show whether or not desired national outcomes are being achieved by the collection of programs addressing HIV/AIDS in the respective country.

The final critical element involves the collaboration of international agencies to determine the global response (e.g., the numbers and types of initiatives aimed at various at-risk populations) to HIV/AIDS and the impact of this response. To understand the collective effectiveness and impact of programs, GAP will need to participate in special impact studies that, as previously defined, draw rigorously collected information from a number of programs.

GAP M&E Framework and Data Types

Figure 16 conveys the logical progression of these program elements (i.e., assessment and planning, inputs, activities, etc.) and types of data associated with these elements. For instance, in the area of program assessment and planning, situation or response analysis and assessment of stakeholder needs may provide the data to assess program needs and planning. In the area of program inputs, information about staff (e.g., numbers of staff), funds (e.g., amount of

funds), etc., provide the necessary data to conduct M&E. In the below framework, data are further organized into categories of program development data collected during the assessment and planning stage, program-based data collected during program implementation, and population-based biological, behavioral, and social data collected when program outcomes and impacts have materialized.

Figure 16.   Evaluation Framework

Global AIDS Program M&E Framework and Illustrative Data Types

Assessment & Planning	Input (Resources)	Activities (Interventions, Services)	Output (Immediate Effects)	Outcomes (Intermediate Effects)	Impact (Long-term Effects)
Situation Analysis	Staff	Trainings	# Staff Trained	Provider Behavior	HIV Incid/Prev
Response Analysis	Funds	Services	# Condoms Distributed	Risk Behavior	Social Norms
Stakeholder Needs	Materials	Education	# Test Kits Distributed	Service Use	STI Incid/Prev
Resource Analysis	Facilities	Treatments	# Clients Served	Behavior	AIDS Morb/Mort
Collaboration plans	Supplies	Interventions	# Tests Conducted	Clinical Outcomes	Economic Impact
				Quality of Life	

Program Development Data

Program-based Data

Population-based Biological, Behavioral & Social Data

In addition to monitoring these illustrative data types, select programs conduct enhanced process and outcome evaluations.

CDC   ORC MACRO

M&E Data and the Public Health Model

The elements and types of data collected for purposes of monitoring and evaluating programs are related to a general public health model that poses questions along the path of understanding diseases and conditions that affect the public. The steps of this model, as well as their related questions, are included in Figure 17. This step-by-step approach, at the most basic level, seeks to identify problems and appropriate questions to track and comprehensively understand public health issues. Once these issues are understood, the next step is to identify the types of interventions that most efficiently and effectively reduce and/or eliminate the problem. Public health professionals have sought to arrive at the best approaches for

assuring that the health of the public is maintained or restored, especially among those most impacted by disease and mitigating circumstances, such as poverty and lack of access to disease prevention and treatment. The illustration below depicts the elements and types of data identified in the GAP M&E framework as they relate to this public health questions approach.

Figure 17. Public Health Questions Approach

A Public Health Questions Approach to Unifying AIDS M&E

Determining Collective Effectiveness	**OUTCOMES & IMPACTS MONITORING**	Are collective efforts being implemented on a large enough scale to impact the epidemic (coverage; impact)? Surveys & Surveillance
Monitoring & Evaluating National Programs	**OUTCOMES**	Are interventions working/making a difference? Outcome Evaluation Studies
	OUTPUTS	Are we implementing the program as planned? Outputs Monitoring
	ACTIVITIES	What are we doing? Are we doing it right? Process Monitoring & Evaluation, Quality Assessments
	INPUTS	What interventions and resources are needed? Needs, Resource, Response Analysis & Input Monitoring
Understanding Potential Responses		What interventions can work (efficacy & effectiveness)? Are we doing the right things? Special studies, Operations res., Formative res. & Research synthesis
		What are the contributing factors? Determinants Research
Problem Identification		What is the problem? Situation Analysis and Surveillance

CDC   ORC MACRO
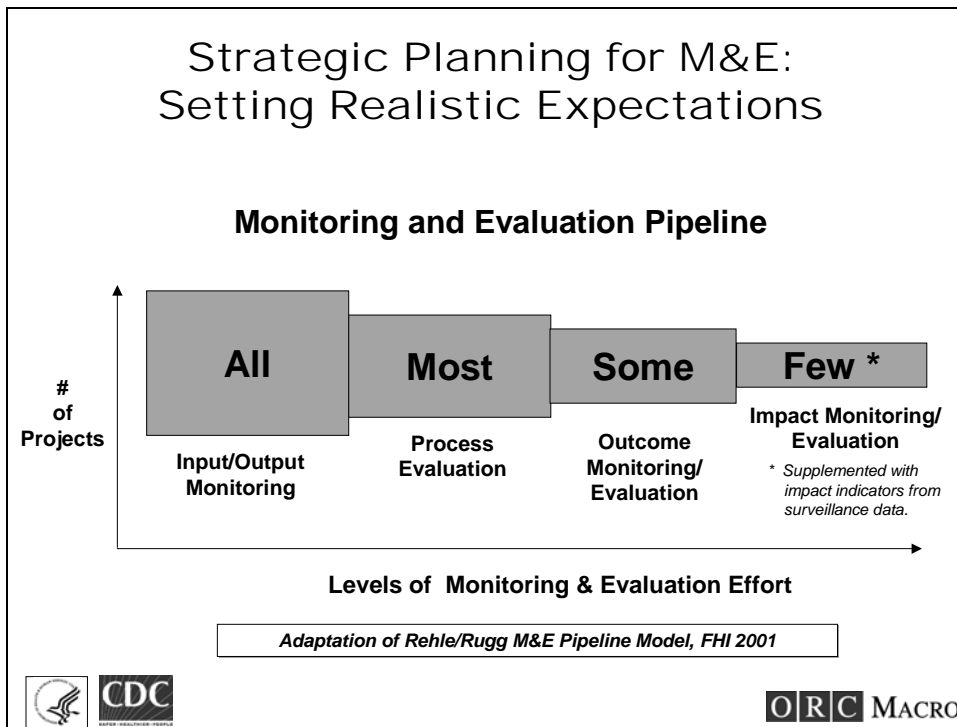
Setting Realistic Expectations

The above approach illustrates the rationale for GAP programs to conduct a number of M&E activities that produce a variety of data. While the possibilities are many, not all programs need to participate in all M&E activities that are part of the larger framework. All programs are expected to participate in the primary levels of M&E, including assessing needs and planning programs as well as monitoring inputs and outputs once implementation begins. Expectations to conduct additional levels of M&E vary by program.

First, programs need to use their resources wisely, so, the extent and costs of M&E activities should be commensurate to the size, reach, and cost of programs. In short, M&E should never compromise or overtake program implementation. Second, not all M&E activities are appropriate for programs or the stages of development at which programs happen to be at a given time. For instance, before conceptualizing and implementing outcome monitoring and evaluation, a fledgling program would need to establish its ability to serve a large enough number of clients with an appropriate intensity of service. It may take several years for a new program to develop in this manner. The illustration of an M&E "pipeline," introduced in an earlier chapter, suggests varying expectations for implementation of types of M&E activities among GAP programs.

Figure 18.   Monitoring & Evaluation Pipeline

Strategic Planning for M&E:
Setting Realistic Expectations

Monitoring and Evaluation Pipeline

# of Projects

All

Most

Some

Few *

Input/Output Monitoring

Process Evaluation

Outcome Monitoring/ Evaluation

Impact Monitoring/ Evaluation

* *Supplemented with impact indicators from surveillance data.*

Levels of Monitoring & Evaluation Effort

Adaptation of Rehle/Rugg M&E Pipeline Model, FHI 2001

CDC

O|R|C Macro®

13. GAP Planning, Budget and Reporting System

> **Important Disclaimer**
>
> This Guide focuses on issues related to monitoring and evaluation for program improvement that can be applied in any setting, and to any program. However, *Chapter 13. GAP Planning, Budget, and Reporting System* and associated Appendices K, L and M sets out GAP-specific planning and reporting requirements for the reporting round completed in 2003. Note that these requirements may be updated for future reporting rounds, hence GAP offices should refer to guidance sent out from the GAP Office of the Director each year to meet their reporting requirements. The issues and templates in this guide are included for illustrative purposes only. Note also that this Field Guide was developed before the implementation of President Bush's Emergency Plan for AIDS Relief, hence it does not include any information on this new Initiative.

Planning and Reporting Mechanisms

GAP has instituted several planning and reporting requirements that are inter-related. Data gathered through M&E activities support the completion and updating of the planning and reporting requirements, which include:

- **The Country Assistance Plan (CAP)**. The CAP consists of planning related to GAP-supported activities in-country. It requires that Annual Objectives are defined by GAP technical strategy in support of the country-specific National HIV/AIDS Program goals. It also includes a listing of planned/ongoing M&E activities for each of the GAP-supported, in-country technical strategies and a full *Monitoring and Evaluation (M&E) Plan.* This is a comprehensive planning document for all M&E activities within GAP Programs. The M&E questions that shape the Plan are directly related to the Annual Objectives stated in the CAP. In other words, the M&E Plan should set forth a strategy for measuring progress toward
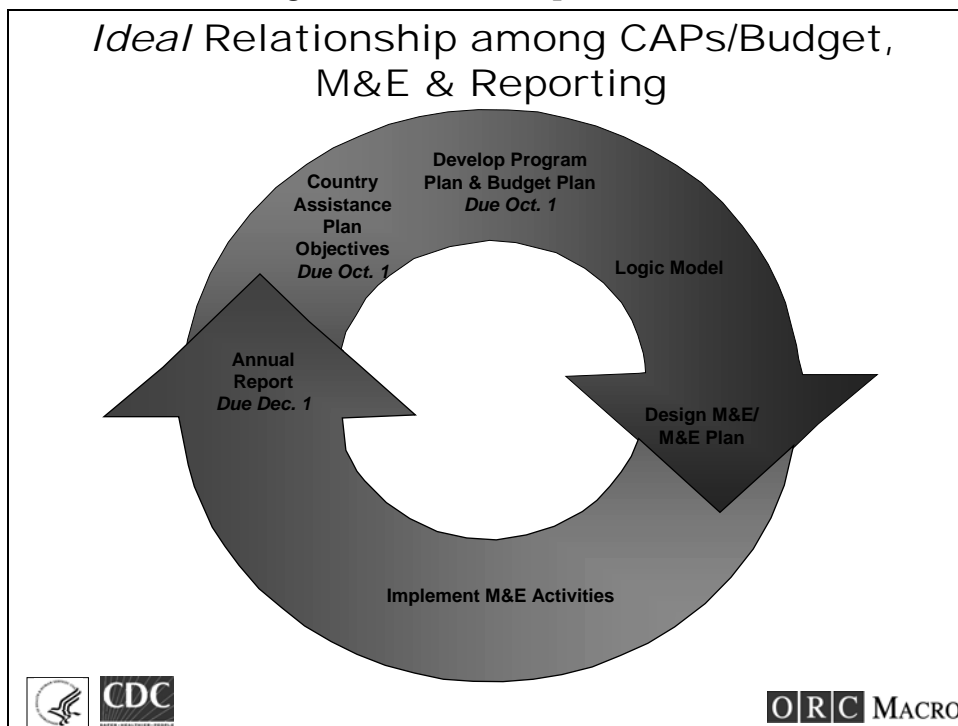
attaining process and outcome objectives relating to each of the GAP-supported technical strategy areas discussed in the CAP.

- **The Budget Plan.** The Budget Plan includes a breakdown of the budget by Object Class, by GAP-supported Projects, and by technical strategy, as well as a listing of contracts over $25,000, and a listing of cooperative agreements GAP is planning to fund/funding in-country.

- **The Annual Report.** The Annual Report provides an opportunity to pause for GAP program managers to relay to GAP HQ their progress, achievements, and challenges in reaching Annual Objectives for the past Financial Year, and their recommendations for next steps. This report also includes data on required indicators by GAP-supported technical strategy areas.

The following graphic shows the way in which these key planning and reporting stages are linked, and specifically, ways in which M&E activities contribute to each.

Figure 19. Relationship Model

Ideal **Relationship among CAPs/Budget, M&E & Reporting**

Country Assistance Plan Objectives *Due Oct. 1*

Develop Program Plan & Budget Plan *Due Oct. 1*

Logic Model

Annual Report *Due Dec. 1*

Design M&E/ M&E Plan

Implement M&E Activities

The upper arrow in the above graphic represents the program planning stage in which CAPs and Annual Objectives are finalized, program and budget plans are developed, corresponding program logic models are designed, and specific and measurable objectives are developed for individual projects or interventions within programs. It is toward the end of this program planning and budgeting phase that the M&E Plan should be developed and finalized. The lower arrow represents program implementation when individual projects and interventions within programs are monitored and, if appropriate, evaluated according to the M&E Plan. Select data arising from M&E activities (i.e., required indicator information and other M&E data that would be pertinent for GAP HQ to know) are then reported in the Annual Report. Information and conclusions from M&E data analysis can also help inform the updating of the CAP and Annual Objectives for the coming year. This progression represents an "ideal" situation; however, variations may occur. For example, monitoring data may be collected and reported before a comprehensive M&E Plan is developed. It may be necessary to report limited M&E data to GAP HQ before the Annual Report is issued [e.g., in response to specific requests from the Office of Management and Budget (OMB)], or, it may be desirable to adjust CAP objectives based on emerging M&E data before the annual reporting round.

The following describes GAP requirements with reference to the CAP, program planning, budget planning, tracking required indicators, and annual reporting.

Country Assistance Plan

The Country Assistance Plan comprises the following information:

1. Demographics and Epidemiology of HIV/AIDS, STIs, and TB
2. National HIV/AIDS Response
3. CDC/GAP Country Profile:
   - Start of Program
   - Staffing
   - Budget
   - Technical Strategies Overview
   - Institutional Arrangements & Collaboration
4. Program Goals, Annual Objectives, and Program Activities
5. M&E Activities
6. Appendix CDC/GAP M&E Plan

See Appendix K for CAP template for FY 04.

The Country Assistance Plan

A systematic planning process for GAP was initiated in 2002 in order to:

- Help country programs and HQ branches focus their activities.
- Provide a programmatic framework for annual budgeting and for assessing progress (through annual reporting and program reviews).
- Facilitate communications regarding GAP programs and priorities both within and outside of GAP.

Plans for GAP programs are formally conveyed in the CAP.

During the first year of the planning process (FY2002), country program managers completed their plans by providing the requested information in a template created in Word. In the future, CAPs will be prepared and submitted through a Web-based or CD/ROM-based (for countries with unreliable Web access) information system.

The Plan template covers two years.[2] Program managers will be asked to review and update Plans on an annual basis. Starting in FY2004, the CAP will also include detailed M&E information. A listing of M&E activities for each of the GAP-supported technical strategies and a comprehensive M&E Plan are required. To assist GAP offices to comply with this requirement, the GAP HQ M&E team will be providing extensive hands-on training and will continue, as in the past, to provide direct TA.

The CAP template is expected to remain similar from year to year, but slight modifications may be needed to meet any potential changes in GAP priorities. Country program managers are responsible for preparing their Plans and submitting them to HQ. The M&E Team serves as a clearinghouse for the Plans and monitors

Template

The template for the FY04 CAP submission which was due October 1, 2003 is provided in Appendix K.

---

[2] In the first year in the planning process, country program managers were asked to provide information for 3 years: retrospective information (i.e., for FY02) as well as prospective information (i.e., for FY03 and FY04).

Plan development. Other HQ staff provide guidance as needed to complete the Plans.

Country program managers circulate their drafts for one round of review and feedback from a minimum of 4 types of reviewers: primary local partner(s), the Country Support Officer, the HQ Technical Lead for the country, and the M&E HQ Point Person for the country. Country program managers are then responsible for incorporating the feedback and finalizing the Plan. CAPs are due to the M&E HQ point person on October 1 of each year.

The Budget Plan

The Budget Plan templates were initiated in 2001. During the first year of the planning process (FY2001), country program managers completed their plans by providing the requested information in a template created in Word and Excel. In the future, program managers will submit budget information through a Web-based or CD/ROM-Based information system.

Budget Plan

The Budget Plan comprises the following information:

1. Budget by object class
2. Budget by projects
3. Budget by technical strategies
4. List of contracts over $25,000
5. List of cooperative agreements

See Appendix L for Budget Plan template for FY04

Program managers will be asked to submit budget information on an annual basis. The budget template is expected to remain similar from year to year, but slight modifications may be needed to meet any potential changes in GAP priorities. Budget Plans are due to the HQ budget team on October 1 of each year.

The Annual Report

The Annual Report, which summarizes overall GAP-supported programs, activities, and achievements, serves a variety of audiences including Congress, OMB, and other external stakeholders; Country Program managers, GAP HQ, CDC OD; and, National AIDS Programs and other implementing partners. It represents the most comprehensive process for:

- Tracking progress of GAP country programs according to stated objectives and documenting major accomplishments and challenges.
- Measuring performance of GAP country programs on a standardized set of program indicators (see "Required Indicators and Data Sources" below).
- Making suggestions for program improvement and future program directions.

Country program managers are responsible for preparing their Annual Reports and submitting them to HQ M&E team by December 1 of each year.   The reporting template is expected to remain similar from year to year, but slight modifications may be needed to meet changes in audience needs. Reports for FY2001 and FY2002 were prepared as Word documents. In the future, reports will be prepared and submitted through a Web-based or CD/ROM-based (for countries with unreliable Web access) information system.

Each year the HQ M&E team, in collaboration with GAP HQ Technical Teams and GAP country programs, will determine if modifications need to be made to the standardized set of GAP program indicators. The selection follows the GAP M&E Framework and is consistent with the UNAIDS Monitoring and Evaluation Reference Group (MERG) guidelines and recommendations, where applicable. The indicators reflect GAP program maturation, i.e., FY2000-FY2002 indicators focus primarily on inputs, activities, and outputs; greater emphasis will be placed on outcomes and impacts in subsequent years. Indicators may also change over time to reflect potential changes in GAP priorities,

---

Annual Report

The Annual Report comprises the following information:

1. Executive Summary:
   - Major Accomplishments
   - Challenges
   - Recommendations for Future Direction
2. Previous FY Achievements & Budget Spent
3. Key Findings from Evaluation Studies
4. Previous FY Indicators

See Appendix M for Annual Report Template for FY 03

Template

The template for the FY03 Annual Report submission which was due December 1, 2003 is provided in Appendix M.

and/or new developments in technical strategies. The current set (i.e., for FY2003 reporting) of standardized indicators is provided in Appendix M with the Annual Report Template for FY2003. Detailed specifications and data-collection strategies for these indicators are provided in a separate manual "Indicator Guide for Annual Reporting."

GAP Planning, Budget and Reporting System

GAP has developed a Web-based data-entry tool for all planning and reporting requirements to HQ. An electronic, systematic approach to enter data greatly improves consistency and quality of information submitted, as well as make data entry and submission more efficient. The GAP Planning, Budget, and Reporting System (PBRS) will also assist HQ in compiling and analyzing the country-specific information in a timely fashion. For example, each year, the HQ M&E team prepares a summary document that synthesizes all country program reports, as well as GAP HQ achievements. PBRS and its output will be used throughout the year by the HQ M&E team, GAP OD, and others to prepare additional briefs and presentations tailored to specific audiences.

14. THE M&E PLAN—OVERVIEW

Each GAP country program is expected to use the guidelines presented in this guide to design and implement a set of M&E activities relevant to its purpose and goals, and to meet HQ requirements for annual reporting. The M&E Plan is a short document that provides a roadmap for how M&E will be conducted over the life of the program; it lays out the means by which program managers will measure changes that occur as a result of the program's objectives and activities.

The Plan should also clearly specify who is responsible for the M&E activities and a timeline for activities and deliverables. This document serves as a management tool to help ensure that M&E functions are designed and carried out in a systematic and timely manner. It also serves as a communication tool to convey to partners and external stakeholders the kinds of information GAP is generating to track program implementation and measure program effectiveness, and to identify opportunities for collaboration.

Each country program is responsible for developing its own M&E Plan. The HQ M&E team offers TA as needed to develop the Plan and will review all Plans to ensure that adequate resources and activities are in place to meet GAP global M&E goals and objectives. Outlined below is a recommended process for developing the Plan.

Establishing the M&E Planning Team Within Each Country Program

Prior to embarking on M&E planning, the GAP program manager should form a team of GAP staff and others, and assign a team leader. Together the team should agree on roles and responsibilities and establish a timeline for drafting, reviewing, and finalizing the Plan.

The development of an M&E Plan for GAP programs, like the design of the program itself, can benefit greatly from participation

---

The M&E Plan should clearly outline data-collection and reporting activities that:

- Focus on key program management issues.
- Focus on the needs and interests of key stakeholders (information users who are both internal and external to the GAP country program).
- Are feasible, practical, and in line with program resources.

Template

Appendix N includes a template for an M&E Plan that you may use as a guide for developing this Plan.

of persons who have a stake in the management and outcomes of the program as well as those who will carry out the M&E activities. Such persons include GAP program managers, program implementers (e.g., grantees, Ministry of Health counterparts), program beneficiaries (e.g., recipients of GAP TA or GAP-supported services), representatives from collaborating funding agencies, and GAP and partner staff with M&E responsibilities. A participatory approach to the design and implementation of M&E activities enhances the likelihood that M&E data are used in program decision-making because key stakeholders determine what information to collect. Additionally, involving those who will carry out the M&E activities leads to the design of feasible methods and systems and greater commitment to the work.

GAP programs can involve their stakeholders in various ways in the M&E planning process. The optimal degree of stakeholder participation will depend on the structure and operations of the program, and on M&E objectives and resources. In some cases, it may be appropriate to form an M&E planning team whose members work together on all the planning steps. In other situations, it may be appropriate to bring in certain types of stakeholders at selected steps in the planning process. At a minimum, a draft Plan should be circulated to key stakeholders for review and comment.

Key Steps to Developing an M&E Plan

Key Steps to Developing an M&E Plan:

1. Align projects and activities with CAP program goals
2. Develop program logic models
3. Identify information users and their needs
4. Frame key M&E questions
5. Determine M&E methods
6. Develop M&E action plans

The below recommended six-step process will help you develop an M&E Plan for each of your technical strategies. Although each step builds on work carried out in the previous step, the first five steps all contribute to the specification of an underlying framework and can be initiated out of sequence and later harmonized. The last step defines the operations necessary to address framework parameters. In practice, steps 1 through 5 may need to be revisited once the operations have been outlined. For example, information users and M&E questions may need to be reprioritized after finding that resources are inadequate or time schedules are unrealistic.

Step 1. Align Projects and Activities with CAP Program Goals

The program goals and objectives of the CAP serve as the foundation for the M&E Plan. The goal in each CAP should clearly state what the program intends to accomplish over a three-to-five-year time period. The accompanying Annual Objectives provide milestone accomplishments toward achievement of these goals and provide the framework for program M&E. For some programs, Annual Objectives may reflect the intended outcomes of programs for each respective year, but there may also be project-level objectives that program staff may wish to develop. These project-level objectives are not necessarily reported in the CAP, but may be very important for program staff to articulate and use as milestones to help track accomplishments and determine next steps. Both types of objectives, the larger Annual Objectives and project-level objectives, should be measurable, meeting the SMART framework.

All program activities should directly support one or more of these objectives and, therefore, support the larger program goal. Typically, these activities will be organized as projects that are either implemented by GAP staff themselves, by Ministry of Health (MOH) counterparts, or by grantees or contractors. If projects or major activities cannot be mapped directly to Annual Objectives and the Program Goal, then program planning should be undertaken to either reframe program objectives and goals or redesign implementation strategies.

Step 2. Develop Program Logic Models

Objectives and activities should be the basis of a logic model that describes the sequence of events for bringing about change in a program or a system of related programs. The logic model combines the main program elements into a picture of how the program is supposed to work. The detailed explanation of the process of creating a program logic model is described in Chapter 5.

Step 3. Identify Information Users and Their Needs

Three factors play a primary role in determining the scope and type M&E activities:

1. Who are the primary stakeholders or users of information about the program?
2. What do they need to know (and for what purpose)?
3. When do they need the information?

If M&E planning involves representatives from key stakeholder groups, then these questions can be answered.

Identification of the M&E needs of users helps to focus M&E activities on the most critical information and also helps to determine the best way to disseminate this information after it is acquired. Making the effort to assess users' needs during the M&E planning phase is a key step to facilitating use of the M&E data and results when they actually become available.

Step 4. Framing the Key M&E Questions

Some data that need to be discussed in the M&E Plan are required monitoring indicator data. Over and above this required information, program staff will want to ask other evaluation questions. Answers to these questions will contribute to decisions for improving programs, as well as provide information for various program stakeholders. These questions will be determined by (1) program goals and program operations; (2) stakeholder needs; and (3) M&E capacity and resources. In some cases, stakeholders may be able to specify the exact questions that they would like to have answered. In other cases, it may be useful to guide them through the process of identifying and prioritizing questions. Logic models can be helpful tools for doing this. They also show the staging or timing of different questions, e.g., those that can be answered on a routine basis and those that can only be answered at fixed points in time or after certain stages of the program have been completed. In most country program situations, M&E questions should be identified separately for each program goal. Refer to prioritization of questions in earlier chapters.
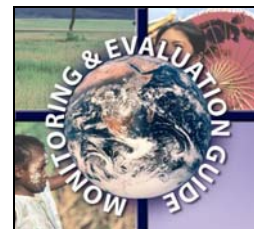
Step 5. Determining M&E Methods

Once M&E questions have been identified and prioritized, it is then necessary to determine the most appropriate data-collection methods for gathering information to answer these questions. Specifically, program staff should identify (1) the types of M&E activities needed to answer the M&E questions (i.e., input/output monitoring, outcome monitoring, process or outcome evaluation); (2) the indicators or measures that will be used/collected (e.g., numbers of clients, amount of funding, etc.); and (3) the data sources for these indicators/measures (e.g., client records, surveys, etc.). For a full discussion on determining methods, see Chapter 8.

Step 6. Develop M&E Action Plans

In most country program situations the M&E planning team should develop an action plan for each program goal (e.g., provide treatment and care, assure blood safety). This action plan should address what action steps will be taken, who will be responsible, what resources will be required, and a timeframe for the following:

- Data collection/validation.
- Data analysis.
- Dissemination of results.

Monitoring M&E Activities

M&E Plans should be used as a general guide for monitoring the progress of M&E activities. This Plan is updated as needed, but at least on an annual basis.

Reasons for the need to update an M&E Plan include one or more of the following:

- Launch of a new project.
- Change in stakeholders.
- Changes in Program logic.
- Need for new or "re-validated" indicators.
- Delay in implementing M&E activities.

An example of an M&E Plan for surveillance from one country field office appears in Appendix O.

Integrated M&E Strategy

All GAP field offices that are addressing more than one technical strategy will also need to integrate M&E plans for each technical strategy into an integrated M&E strategy. The latter may include explanations of ways that M&E resources (e.g., staff, financial, materials) may be shared across technical strategies; ways in which similar data collection methods and data themselves may be used to evaluate two or more strategies at the same time; and other areas of integration of M&E activities. The M&E strategy may be submitted as a narrative summary. Together with the M&E plans for individual technical strategies, this integrated strategy forms the overall M&E plan for the GAP program.