

PH - 70-2-015
Jan 6 1997

MODULE



1

Acknowledgements

The following AIDSTECH and AIDSCAP staff contributed to the content and production of this document (in alphabetical order);

Patricia Bailey
Susan Hassig
Jan Hogle
Michele T. Villinski
Sharon Weir
Leslie Young

Graphic Design

Sandy Paul

MODULE



1

Table of Contents

Section

| | | |
|----|--|----|
| 1 | Rationale | 1 |
| 2 | Definitions: evaluating AIDS prevention programs | 3 |
| 3 | AIDSCAP evaluation strategy | 5 |
| | Country level and subproject level | 5 |
| | Types of evaluation | 6 |
| | Baseline assessments | 6 |
| | Priority prevention indicators (PPIs) | 7 |
| | Management information system (MIS) | 7 |
| 4 | A note on project design | 9 |
| 5 | Outcome and impact evaluation | 11 |
| | Methods | 12 |
| | Comparative or attainment? | 12 |
| | Impact indicators: HIV and STD incidence | 14 |
| | Additional sets of outcome and impact indicators | 15 |
| 6 | Formative evaluation | 17 |
| 7 | Process evaluation | 19 |
| | Examples of process indicators | 20 |
| 8 | Reliability and validity | 23 |
| 9 | Use of multiple evaluation methods | 25 |
| | Qualitative and quantitative complementary methods | 25 |
| | Methods for assessing change | 27 |
| | Examples of the use of multiple methods | 29 |
| 10 | Appendices | 31 |
| | References and bibliography | 31 |
| | WHO/GPA's Priority Prevention Indicators (PPIs) | 34 |
| | AIDSCAP Project Logical Framework Summary | 37 |

SECTION



1

Rationale

The idea for an evaluation tools guide was originally conceived by AIDSTECH evaluation staff and published in 1992 as a single volume called "Tools for Project Evaluation: A Guide for Evaluating AIDS Prevention Interventions". AIDSCAP is expanding on this practical approach to evaluation with a series of modules designed as references for project managers, designers, and field staff. As such, the emphasis throughout is on practical application of evaluation theories. The series of modules was designed with the hope that better understanding of the purposes and methods of evaluation would lead to more creative and thorough implementation of evaluation components.

The objectives of this series include the following points:

- To orient AIDSCAP implementing agencies and resident advisors to the project evaluation strategic plan
- To provide practical, not theoretical, step-by-step, methodological guidelines for project implementors conducting evaluation
- To provide separate, small volumes, each with a specific methods focus, in a user-friendly, concise format, with tables and diagrams
- To publish one volume at a time so as to keep recipients alert to evaluation issues periodically
- To provide a standardized and concrete approach for HIV/AIDS/STD prevention evaluation activities

Throughout the life of the AIDSCAP project, the modules series will continue to be published. The Evaluation Unit intends to reissue various modules, updated to include concrete examples from AIDSCAP interventions worldwide. This series will thus serve as a source for evaluation "lessons learned".



Definitions: evaluating AIDS prevention programs

“Evaluation” is a concept that means different things in different settings. In standard English, the term means “to determine the value of”, “to examine carefully”, “to appraise”. Evaluation is perhaps commonly thought of as something that is done at the end of a project to determine whether or not the project’s objectives have been achieved. However, as it is currently conceived by the AIDSCAP project, evaluation is a more comprehensive construct implying a solid foundation in the design phase of AIDSCAP interventions. Evaluation of AIDSCAP interventions began with the writing of the cooperative agreement between the United States Agency for International Development (USAID) and Family Health International (FHI) in which the strategic importance of evaluation was first emphasized. AIDSCAP’s methods of conducting evaluation have been further refined by the AIDSCAP Evaluation Strategy, developed early in 1992 and included in the AIDSCAP Technical Strategies document.

A good working definition for “evaluation” was proposed by the Committee on AIDS Research and the Behavioral, Social, and Statistical Sciences which defined the term with the following series of questions:

“Evaluation is a systematic process that produces a trustworthy account of what was attempted and why; through the examination of results - the outcomes of intervention programs - it answers the questions, “What was done?” “To whom, and how?” and “What outcomes were observed?” Well-designed evaluation permits us to draw inferences from the data and address the difficult question: “What do the outcomes mean?” (Turner, Miller, and Moses 1989:317-318, cited in Coyle et al 1991:16)

AIDSCAP operationalizes this definition further with an emphasis on solid evaluation design *prior to* implementation of interventions. Project evaluation, as conceived by AIDSCAP, should be focused, practical, easily implemented, and yield information about the progress of an intervention toward its objectives.

From AIDSCAP’s perspective, the goal of evaluation is to obtain information to enable policy makers, implementing staff, and managers to measure the effectiveness of strategies and interventions to prevent the sexual transmission of HIV. Ideally,

managers, technical assistants and researchers work together to develop evaluation methodologies that maximize the quality of information acquired and minimize the burden on field staff. Ultimately, evaluation results are used to improve project implementation and inform new project design.

Evaluation, as treated in this series of modules, does not investigate questions better answered by large-scale social science research.

Although the scope of the evaluation may be limited by funding, personnel, or other resources, some means for assessing the success of any project or activity must be devised. An analysis including only management or process indicators lacks the generalizability and interesting comparisons that can result from more extensive evaluation, but can give some indication of the project's progress toward its objectives.



AIDSCAP evaluation strategy

The AIDSCAP project's overall goal is to reduce the rate of sexually transmitted HIV infection in selected priority countries in Africa, Asia and the Latin America/Caribbean regions. In order to accomplish this goal, countries must implement interventions most likely to alter those behaviors known to facilitate transmission of the virus. Limited resources must be divided among the three primary technical strategic approaches — STD diagnosis and treatment, condom distribution and promotion, and communication for behavior change—and the supporting strategies of policy development, behavioral research and evaluation. AIDSCAP's approach to evaluation can be summarized as a three-dimensional matrix of:

- level (country program and subproject)
- type (formative, process, outcome and impact)
- method (qualitative and quantitative).

Country level and subproject level

Evaluation will occur at the global level of the “country program” as well as at the level of the numerous subprojects which make up each country program. Even in countries where AIDSCAP is not involved in nationwide program implementation (such as Thailand, Brazil and Nigeria), the summary outcome and impact of the program components will be monitored and evaluated at the aggregate or global level.

Each sub-project within the country program will be designed in cooperation with locally-identified NGOs, PVOs, and AIDSCAP staff. Each sub-project proposal contains an evaluation plan incorporating those country program indicators which are relevant to the particular sub-project, along with process and outcome indicators specific to the sub-project.

Sub-project evaluations will often provide interim estimates of progress to objectives for the country program as a whole, in addition to their primary purpose of measuring the success of a particular sub-project.

Types of evaluation

At both levels, baseline assessment (including formative research) measures key indicators which are compared to subsequent follow-up (outcome) measurements. A set of key, country-level process indicators is regularly measured, usually each month. At the subproject level, process indicators appropriate for the particular subproject will also be tracked and aggregated as needed for country-level measurements.

AIDSCAP considers impact indicators to be the biological measures of HIV and syphilis seropositivity in antenatal clinic populations of women aged 15 to 24, and generally tries to use existing sentinel surveillance data rather than support original data collection. In countries without sentinel surveillance systems, or where existing systems need upgrading, AIDSCAP provides technical assistance to improve the quality of information obtained.

Each type of evaluation at both levels (country program and subproject) features complementary quantitative and qualitative methods for measuring and interpreting key indicators. Using both methods ensures the most comprehensive understanding of the complex and sensitive issues facing HIV/AIDS prevention programming. At the country program level, more emphasis is placed on behavioral outcome indicators, while at the subproject level, process indicators measuring project outputs are emphasized. AIDSCAP also is currently developing protocols and instruments for measuring capacity building of the PVOs, NGOs and government institutions with whom AIDSCAP works.

Baseline assessments

Comprehensive baseline assessments of AIDSCAP country programs are conducted at the beginning of the implementation phase. These assessments use existing data wherever possible but also might include some degree of formative research such as a population-based survey focusing on knowledge, attitudes, beliefs and practices, complemented by focus group discussions.

Baseline assessments include several components reflecting AIDSCAP's comprehensive strategies for reducing sexual transmission of HIV. These components include:

- a review of ***HIV and STD prevalence*** in the nation and in targeted groups and regions, along with an assessment of the methodology of collecting blood samples (the sentinel surveillance system)
- an analysis of the supply, distribution and promotion of ***condoms*** in the nation and the intervention regions
- an assessment of the status of ***STD services delivery*** in health facilities nationwide and in the intervention regions with an emphasis on the role of the biomedical system in the health seeking behaviors of target populations
- a brief description of significant ***behavior change communication*** strategies, campaigns, messages, materials, training efforts, curricula, and targeted populations, country wide and specifically in intervention regions
- a description of the ***policy environment*** for HIV/AIDS interventions nation-

wide, including the constraints facing interventions and possible strategies for changing the policy environment

- **baseline quantitative and qualitative behavioral data** usually collected via KAPB surveys complemented by focus group discussions
- a literature review of **social and behavioral science research data** on sexual behavior relating to HIV/AIDS available in-country and through AIDSCAP library resources
- **institutional assessment** to analyze organizational capabilities, and training and technical assistance needs

AIDSCAP has developed a set of core KAPB (knowledge, attitudes, beliefs and practices) questionnaires adapted for the specific types of populations accessed by AIDSCAP-supported interventions, including female sex workers, male clients of female sex workers, men who have sex with men, youth, and the general population. There are also focus group discussion guides for youth. These instruments can be adapted for use in each country program for baseline and follow-up surveys. Data from these surveys, complemented by qualitative data, provide comprehensive (that is, generalizable and in-depth) information on knowledge, attitudes, beliefs and practices of at-risk populations.

Priority Prevention Indicators (PPIs)

AIDSCAP works closely with WHO's Global Programme on AIDS (GPA), the Centers for Disease Control (CDC), the United States Agency for International Development (USAID), and other agencies on the continuing development of Priority Prevention Indicators (PPIs) for national AIDS control programming. This minimum set will enable cross-national comparisons, as well as guide the development of larger sets of country and subproject indicators.

These PPIs are not intended to provide all the information necessary to monitor the implementation of prevention programs, nor to measure the impact of all prevention and control activities. Therefore, AIDSCAP evaluation plans also include additional indicators depending on the specific components (subprojects) of country programs. These additional indicators may be developed during implementation planning or during development of subprojects.

Within each country program, AIDSCAP retains GPA's wording of the PPIs in all survey instruments, and will use GPA protocols in assessing condom acceptability and accessibility, and STD case management. Copies of these protocols and instruments can be obtained from AIDSCAP headquarters or regional offices, or from WHO/GPA.

Management Information System (MIS)

The AIDSCAP Management Information System (MIS) facilitates the flow of information within the AIDSCAP project, enabling AIDSCAP to efficiently apply "lessons learned" at the earliest possible opportunity, continually assess progress toward objec-

tives and goal, and respond efficiently to requests for information from organizations and individuals.

The MIS functions as a central point for coordinating evaluation data collected by subprojects and for country program evaluation. The MIS is linked to AIDSCAP regional offices and their computer facilities, and provides country offices with fast feedback and inter-region comparisons.



A note on project design

IHIV/AIDS/STD prevention programs are difficult to implement under any circumstances. Program evaluation should improve interventions and not impede their implementation or discourage participation by members of the target populations. Planning an integrated evaluation process requires flexibility, and an understanding of the pressures on and purposes of various actors in the evaluation design and implementation process.

Particularly in the design phase, tensions can potentially exist between program implementors, on the one hand, and epidemiologists and evaluators on the other, and should be recognized. These tensions arise from the program manager's need to implement a flexible and effective project with limited resources, and the evaluator's desire for reliable, valid and well-planned data collection. Recognition and understanding of these tensions should help interventionists and evaluators be aware of the other's perspective and facilitate cooperative planning prior to implementation.

Tensions in Evaluation Planning

| | | |
|------------------------------|-----|-----------------------------|
| Iterative/ongoing evaluation | vs. | Pre-post summative |
| Adaptive programming | vs. | Controlled study conditions |
| Use of program data | vs. | Unique data collection |
| Phased outreach to all | vs. | Nonintervention sites |
| Sustainable programs | vs. | Field intervention trials |
| Local management | vs. | Outside experts |
| Urgency and flexibility | vs. | Planned timetables |

Planning a project and designing an evaluation plan are inseparable activities. In order for evaluation to be useful to program staff, and to improve implementation, implementing agencies and their technical assistants cannot postpone evaluation planning until the final stages of the project. Instead, planning for evaluation of AIDSCAP interventions begins in the program design phase prior to implementation. To facilitate evaluation activities both during and after implementation (thus reducing tension between implementors and evaluators), project designers should:

- define an initial target population and potential subsequent target populations including geographic location
- estimate the size of the target population as best possible
- detail project objectives – **goal, purpose, outputs, activities** – within a logical framework where each objective includes only one desired result
- develop **objectively verifiable indicators** of success linked to project objectives and activities
- determine what **means of verification** will be used to identify the indicators
- establish a budget and schedule for meeting project objectives
- collaborate with project evaluators to develop mutually agreed upon indicators, and a monitoring and evaluation timeline

The “logical framework” (or “logframe”) technique is a particularly useful method for simultaneously designing projects and their evaluation plans. The logframe process specifies very clearly the sets of indicators needed to conduct an evaluation, and the means of verification for each indicator (ie, how the data for the indicators will be obtained). The technique also accomplishes the goal of allowing the design of a project to emerge from the negotiations of a group of people developing the contents of the four-by-four matrix that summarizes the project objectives. Developing a logframe is a process which is most productive when it is facilitated by someone who knows the technique well and has experience working with groups of people to develop project designs.

AIDSCAP subprojects are developed using the logframe approach. Each subagreement contains a logframe summary of the project. Evaluation plans are thus relatively easy to design at both the country program level and the subproject level, because the logframe contains lists of indicators attached to each objective in the hierarchy (goal, purpose, outputs, activities).

AIDSCAP provides technical assistance in project design and logframe development as needed by program managers, implementing agencies, NACPs, regional offices, and resident advisors.

Additional considerations in evaluation design, which should also be reflected in the project budget, include the following:

- deciding which populations to survey at baseline and follow-up, and establishing a comparison group
- assessing project sources for evaluation data (some potential examples include financial reports, condom audits, HIV/STD testing, IEC pretesting, clinic attendance forms and reports, medical records, staff time sheets, AIDSCAP resident advisor monthly reports)
- assessing non-project data sources (some potential examples include KABP or other regional or national surveys, HIV and STD sentinel surveillance, WHO/GPA AIDS surveys, Demographic and Health Surveys, National AIDS committee data and reports)



Outcome and impact evaluation

Outcome and impact evaluation identify the effects or consequences of a program and ideally attempt to establish that the observed changes are indeed the result of the intervention. These measurements can be made at the end of a project or intermittently during the project, and seek to answer the questions,

“What changes were observed and what do they mean?”

and

“Do the interventions make a difference?”

In general, the objectives of outcome evaluation of interventions to slow sexual transmission of HIV are to identify the program’s impact on:

- prevalence (or incidence) of HIV or STD infection in the target population
- sexual behaviors—condom use, number of non-regular partners
- knowledge and attitudes—of ways to prevent HIV transmission, and self-perception of risk, for example

The success of a program may be related to variables that are difficult to measure in surveys, such as quality of interaction between the high-risk community and health providers. Thus, quantitative outcome results are enriched by accompanying qualitative information derived from methods such as in-depth, semi-structured interviews, participant observation, focus group discussions, and focused (rapid) ethnographic studies.

For example, more subtle aspects of the quality of STD services—such as the extent to which services are accessible, confidential and private, the waiting times minimal, and the staff well-trained—are best evaluated qualitatively through participant observation. Researchers equipped with checklists could be sent to STD clinics to record their observations. After a few days of observation, well-trained observers can produce accurate and consistent assessments.

Methods

The most commonly used methodologies for outcome evaluation are cross-sectional surveys, prospective or cohort studies, and sentinel surveillance.

Cross-sectional surveys among the intervention group (and a comparison group if available) provide information about the prevalence of knowledge, beliefs and types of behavior at a given point in time. A cross-sectional survey conducted only after program implementation documents the existing level of risk behavior post-intervention. Since it does not provide information *prior to* the intervention for comparison, a single survey cannot provide convincing evidence of program effectiveness.

Prospective or cohort studies enable investigators to measure change at the individual level over time. Repeated surveys (e.g., before and after studies) can monitor changes in target population behavior over time. While more persuasive than one-time studies, pre- and post- studies still cannot attribute observed changes to the intervention with certainty, especially if the samples are different. Surveys conducted among target group members before and after the project do not answer the question, “What would have happened to project participants if they had not been exposed to the intervention?”.

Repeated surveys of the same people provide more reliable information than two cross-sectional surveys which do not strictly control who is being interviewed. However, many of the target populations of HIV/AIDS interventions are highly mobile and interviewing the same cohort of individuals months or years later may be nearly impossible. Prospective studies require a high response rate at the follow-up interview. Achieving high response rates usually implies high survey costs. If contacting the same people months later seems highly unlikely, two cross-sectional surveys are recommended.

Ideally, outcome evaluation research should collect pre- and post- data from members of the target community who are randomly divided into an intervention and a control group. This method would provide baseline and follow-up data for individuals within the scope of the project and for individuals who were not exposed to the intervention. These data would be the basis for conclusions about the project’s impact, and would control for changes in behavior due to factors outside the project. Practically, this kind of randomized experiment is seldom feasible. It requires an enormous amount of data collection and is difficult to implement in highly mobile populations, in addition to posing ethical questions.

Selection of outcome evaluation methodologies for a particular intervention depends on the intervention’s objectives, the stage of implementation, and available funds, personnel and analytic expertise. Within these constraints, program staff must design the strongest possible evaluation.

Comparative or attainment?

Two types of indicators are useful in outcome evaluation: comparative and attainment indicators. **Comparative indicators** use data collected more than one time during the intervention to assess changes in an outcome of interest, and require some estimate

of baseline status in order to set a target for accomplishment by the end of the project. **Attainment indicators** measure the status of outcome indicators at a specific point in time, and are particularly useful when “pre” data are unavailable.

AIDSCAP projects will always seek to collect baseline data during the early stages of interventions. However, when designing a project, prior to the awarding of funding via subagreements, measurable indicators of project accomplishments have to be agreed upon between AIDSCAP and implementing agencies. These indicators can be worded as either comparison or attainment indicators. It may be necessary, after baseline assessment is completed, to adjust the project’s targets.

Examples of both comparison and attainment indicators defining the same target are presented in the table.

| | COMPARISON INDICATORS | ATTAINMENT INDICATORS |
|------------------------------------|--|--|
| | Assessed two or more times during the project in the target group and a comparison group if possible | Assessed one time during or after implementation in the target group and a comparison group if possible |
| CONDOM USE | 100% increase in self-reported condom use with non-regular partners [Baseline data indicate that only 30% currently report condom use with non-regular partners; 100% increase would result in 60% reported condom use] | 60% of population surveyed report condom use with non-regular partners |
| | 50% increase in proportion of CSWs reporting use of condom during the last sex act [If pre=60%, then post =90%] | 90% of CSWs surveyed report using a condom during the last sex act |
| | 100% increase in proportion of male STD clients who have ever used a condom [If pre=40%, then post=80%] | 80% of male STD clients report having used a condom at least once previously |
| | 50% increase in proportion of people surveyed who spontaneously report condom use when asked about their AIDS prevention behavior [If pre=50%, then post=75%] | 75% of population surveyed spontaneously report condom use when asked about their AIDS prevention behavior |
| USE OF STD DIAGNOSIS AND TREATMENT | 100% increase in proportion of population surveyed who can name one high-quality source of care for STD symptoms [If baseline=15%, then follow-up=30%] | 30% of population surveyed can name at least one high-quality source of care for STD symptoms |
| | 100% increase in proportion of STD clinic patients who receive a condom and advice on partner notification at the time of treatment [If pre=40%, then post=80%] | 80% of STD clinic patients receive a condom and advice on partner notification at time of treatment |
| SEXUAL PARTNERS | 50% decrease in self-reported mean or median number of different sex partners of army recruits in the last 6 months [If pre=5, then post=2.5] | Mean or median number of different sex partners in the last 6 months reduced to 2.5 among army recruits |
| | 50% decrease in proportions of men reporting 3+ partners in last 6 months [If pre=10%, then post =5%] | 5% of men report 3+ partners in last 6 months |

Impact indicators: HIV and STD incidence

Biological indicators should be carefully chosen. The most persuasive epidemiologic impact indicator for HIV/AIDS prevention programs is a demonstrable increase in HIV incidence that can be linked to the program. While convincing, HIV incidence is also costly and time-consuming to obtain, relative to STD incidence. Ideally HIV incidence data would be obtained by monitoring two cohorts from the same target group. One cohort would be exposed to the intervention, the other cohort (the control group) would not. The evaluation would compare the difference over time in the two cohorts' seroconversion rates. This ideal method is not practical, however, because it would require a massive data collection effort, incredible resource commitments, and be quite difficult to implement in a highly mobile population.

Other difficulties with using HIV incidence as an indicator are: ethical concerns regarding withholding potentially successful interventions from the control group, and problems with measuring changes in incidence where initial rates of HIV infection are extremely high or extremely low.

Sentinel surveillance traditionally monitors HIV or STD infection among convenient samples nonrandomly "representing" the target group at regular intervals over time. The AIDSCAP project considers HIV and syphilis seroprevalence to be the key impact indicators of intervention success at the country program level. Changes in HIV seroprevalence rates, however, are unlikely to be identifiable over the short or even medium term. Changes in STD rates can decrease rapidly over the short term and are used as a proxy marker for changes in HIV incidence. They are also good goal-level indicators for subprojects where support is generally provided over a shorter term.

The advantages of using STD incidence or prevalence as an indicator include:

- If incidence rates of STDs are relatively high, then changes in the rate may be measurable
- Some STDs are symptomatic in men
- If a person has a non-viral STD, then changes in their behavior will be reflected in their STD status
- The frequency of new bacterial STDs may be a useful indicator of behavior change.
- There are fewer ethical concerns about confidentiality and informed consent
- STD surveillance may be the only biological indicator to monitor program impact in populations with very low or very high HIV prevalence or very low HIV incidence

For individual, relatively short term projects, a project design team should decide which STDs to monitor, and whether or not to use indicators requiring laboratory testing. There is a wide range of variation in the costs and complexities of testing for the various STDs commonly found in AIDSCAP intervention regions.

Techniques of measuring STD incidence or prevalence differ greatly between men and women. Women are usually asymptomatic when first infected with an STD, and often do not seek medical attention until complications develop. Physical and laboratory examinations are always necessary for diagnosis. Specific tests to be done depend on local laboratory capabilities, and range from syphilis serology (gynecologic exam not necessary) to tests and cultures that require speculum exams, microscopes and reagents.

Among groups of high risk women, such as commercial sex workers, dramatic changes in STD prevalence can be seen if:

- they decide as a group to adopt the use of condoms with all of their partners
- their solidarity is strong enough to hold them to their decision, even in the face of client resistance.

STD surveillance of men is a useful marker for behavior since the behavior changes necessary to stop the spread of HIV and other STDs depends largely on male behaviors. In many cultures, men choose how many sexual partners they will have, and whether they will use condoms. Self-reported STD history is not a reliable methodology for collecting these data, however.

Additional sets of outcome and impact indicators

In addition to the biological impact indicators, there are various sets of outcome indicators that can be measured in HIV/AIDS intervention programs. Outcome indicators for an ***STD diagnosis and treatment program*** might include:

- increased proportion of patients intending implementation of risk reduction behaviors (eg. condom use, reduced numbers of partners)
- decreased proportion of clients who are repeat attenders at STD clinics
- high proportion of counselors who exhibit and maintain satisfactory skill levels at assessment and supervisory visits.

Outcome indicators for an ***HIV prevention in the workplace program*** could be assessed through an experimental design measuring the following survey- and record-derived elements in work sites which are and are not participating in the supported activities:

- knowledge of HIV prevention methods
- reported condom use, by age group and type of partner
- reported number of sexual partners in a defined timeframe, by age group and marital status
- history of STDs from clinic records (and reported STD history, if method has been validated).

Outcome indicators for an ***HIV counseling and testing program*** might include:

- self-reported intended behavior change (recorded after counseling - eg. intended condom use, reduction in numbers of partners)
- self-reported re-test rate for HIV- and HIV+ clients, extracted from testing center log
- sample of clients recruited at 24, 36 and 48 months to participate in group discussions and/or in-depth interviews regarding their reaction to the services offered at the center(s)
- semiannual gender and age-specific HIV+ rates at center and from sentinel surveillance

Outcome and impact indicators for a ***condom social marketing program*** might include:

- reported condom use, pregnancy and STD history, from a survey of three urban centers undertaken at the end of year X, to include questions to assess awareness of the CSM campaign and recognition and use of the CSM product
- at the end of X years, the infrastructure of the indigenous condom distributor should be able to maintain the condom social marketing activity without donor involvement in actual implementation

Outcome and impact indicators for ***capacity building and institutional development*** activities might include:

- percentage of organization staff who are more confident in their organization's capability to manage HIV/AIDS interventions
- percentage of organizations involved in national level strategic planning for HIV/AIDS interventions
- percentage of organizations able to sustain activities independent of direct AIDSCAP financial and technical assistance support.

SECTION



6

Formative evaluation

Formative evaluation research is used during project design, during baseline assessments, and in the early stages of implementation to:

- evaluate the need for a new intervention
- provide first-hand experience with the target population
- establish baseline measures of key indicators
- determine the best implementation strategy
- stimulate and explore new strategy ideas

Formative evaluation is done just prior to or during baseline data collection, and can include 'needs assessments' when the nature of the intervention requires detailing the specific perceived needs of a target population.

In HIV/AIDS prevention programs, formative evaluation can help to:

- reveal attitudes of the target population towards AIDS, HIV infection, STDs, and prevention options
- develop and pretest educational messages
- assess existing condom distribution systems (including storage facilities, wholesale distribution, and retail distribution)
- develop and pretest condom promotion and packaging
- survey target group preferences for condom distribution alternatives
- identify patterns of health care services utilization that often include use of traditional or alternative resources
- identify explanatory models of sexually transmitted disease causation, symptoms, and traditional treatment protocols, including indigenous language disease names

- obtain information for questionnaire design (e.g., wording of sexual terms; indigenous language categories of types of sexual partners)
- identify, pretest and improve questionnaire items
- provide qualitative data for interpreting quantitative data results
- identify problems early in program implementation

Methods of data collection used in formative research include both quantitative and qualitative techniques. Both types of methods should be used within the same populations for a complete, generalizable and in-depth understanding of the issues involved.

SECTION



Process evaluation

Process evaluation addresses two broad questions:

“What has the intervention done since its implementation?”

and

“Whom has the intervention reached and how has it reached them?”

Process evaluation:

- does *not* indicate if the program has changed the knowledge, attitudes or behaviors of the program’s participants
- examines trends over time to see if an intervention is accomplishing its initial objectives
- should occur throughout the course of a project and at specific time intervals to ensure that the components of a project are being delivered to the right people in an appropriate fashion
- can guide modification of the intervention by identifying both strong and weak areas
- help to make decisions about the management of the intervention, including where additional technical assistance may be needed
- can be used to identify any necessary changes in the indicators used to measure project outcome
- can be used to explain the results of outcome evaluation, clarifying why the program “worked” or why it did not

For example, an education and condom promotion program may “fail”, (e.g., show low reported condom usage or continued high HIV prevalence or incidence) for any of the following reasons:

- inadequate storage facilities leading to substandard physical integrity of condoms
- poor distribution of condoms by outreach workers
- good distribution but ineffective education on condom use

Research designs, or means of verification, used in process evaluation include:

- direct observation of intervention components by program monitors
- focus group discussions,
- in-depth interviews and surveys of program staff and participants
- record keeping, such as the completion of process indicator forms (PIFs)

Focus group discussions can provide an accurate picture of the intervention from the perspectives of the provider and the recipients. Interviews or limited surveys with members of the target population help determine what components of the intervention are being delivered to which groups.

Keeping records of project activities (e.g., condoms distributed) facilitates a true account of what is actually being done throughout the project. Standardized reports (PIFs) aid record keeping by facilitating consistency (over time and across projects) in reporting indicators.

Examples of process indicators

Project-specific process indicators grow out of the scope and shape of the intervention itself. An intervention that attempts to control HIV through improvements in treatment and prevention of STDs will have different process indicators than an intervention based on AIDS education and condom promotion in the general population.

Process indicators to track an ***STD case management*** intervention might include:

- number of clients tested for STDs by gender, age, type of test
- number of clients making at least one repeat visit to clinics, by gender, age, diagnosis
- number of clients clinically examined and found positive for symptoms of STDs, by gender, age, diagnosis
- number of clients receiving laboratory tests and found positive, by gender, age, type of test
- number of STD patients counseled about STD and HIV prevention
- number of STD patients counseled about partner notification
- number of STD patients receiving HIV testing and counseling
- number of STD patients receiving condoms
- number of condoms distributed

- number of printed educational materials produced and distributed
- number of persons contacted through outreach activities

STD facility management process indicators to be tracked periodically might include the following:

- number of clients attending clinics, by gender, age
- number of providers trained in proper diagnosis and treatment of STD, by type
- number of provider training sessions held
- number of condoms distributed/sold through STD treatment intervention sites
- number of counselors trained appropriately
- number of counselor training sessions held
- number of STD diagnostic and treatment facilities “upgraded”
- number of participants at training sessions for lab personnel/educators/clinicians or other providers
- amount (days/weeks) of technical assistance (by technical area) provided to a country program’s STD component

Process indicators for a **condom promotion and distribution program**, on the other hand, might include:

- number of sites visited
- number of outlets (by type) distributing condoms
- number of condoms distributed or sold, by type of outlet and region, on a monthly basis
- number of people contacted at the sites
- number of participants at training sessions for peer educators
- number and types of promotional materials distributed, and number of sites receiving materials
- intercept studies to determine who is using the product, under what circumstances and with what frequency, to allow for improvement of promotion and distribution efforts
- frequency of stock-outs

Process indicators for a project on **HIV prevention in the workplace** might include:

- number of employees and management personnel participating in HIV prevention activities

- number of condoms distributed or sold, by type of outlet and region, on a monthly basis
- number of work sites participating in HIV prevention activities, and allowing such activities on company time
- number of work sites contributing financially to HIV prevention activity (and amount of contribution)
- number of workshops and training sessions held for industry and union leaders, and number of individuals participating.

Process indicators for an ***HIV counseling and testing program*** might include:

- number of individual clients/couples receiving counseling and testing
- proportion who do not return for final test results
- proportion of counseling person-days provided by volunteers
- number/type of training activities conducted
- turnover rate for staff and volunteer counselors
- number of condoms distributed or sold

Process indicators for ***capacity building or institutional development*** activities might include:

- number of management training sessions conducted
- number of staff trained
- number of collaborative relationships developed

Collecting data on process indicators is usually easily accomplished by fulfilling project reporting requirements on a monthly or quarterly basis. AIDSCAP uses a monthly process indicator form (PIF) to facilitate process evaluation of AIDS interventions. The PIF records monthly activity in a quantitative indicator format, and allows for additional, qualitative data to be obtained through open-ended questions.

SECTION



8

Reliability and validity

In both quantitative and qualitative research, it is critically important to use methods and variables that are reliable and valid. Reliability and validity are used as indicators of the research's objectivity, and lend credence to the results.

Reliability is the degree to which the research findings are independent of accidental circumstances of the research, or "the degree of stability exhibited when a measurement is repeated under identical conditions. Reliability refers to the degree to which the results obtained by a measurement procedure can be replicated" (Last 1983). A reliable method is an accurate, consistent and stable measuring instrument; it will yield the same answer however and whenever the procedure is carried out. Reliability in qualitative research depends partly on explicitly described observational procedures.

Measurement validity is "an expression of the degree to which a measurement measures what it purports to measure" ...Study validity is "the degree to which the inference drawn from a study, especially generalizations extending beyond the study sample, are warranted when account is taken of the study methods, the representativeness of the study sample, and the nature of the population from which it is drawn" (Last 1983). In other words, validity is the degree to which the research finding is interpreted in a correct way. A valid method or variable really measures what it claims, and is relevant. Validity is important to consider when naming variables and when deciding which questions to ask and how to word them.

These concerns about reliability and validity are key for behavior change indicators such as change in condom use, change in reported number of partners, change in sexual practices, change in partner mix, change in perceived risk, increase in monogamy, and increase in age at first coitus. For example, it is critically important to obtain valid, reliable estimates of changes in condom use if these estimates are used to assess the effectiveness of an intervention. Unfortunately, obtaining such estimates has not been easy.

High-risk behavior group members who have been educated about the need to use condoms may be willing to report regular use of condoms, whether or not they actually

use them regularly. This form of biased response to survey questions can be alleviated to some degree by the use of multiple methods of data collection. For example, there are a variety of methods for measuring condom use since data on self-reported condom use is often inaccurate. The following methods can increase the validity of condom use data:

- survey questionnaires with multiple questions or methods of asking about self-reported condom use
- comparison of self-reported condom use with a more objective source, such as peer health educators' reports of the individual's condom use or condom distribution records
- use of focus group discussions to obtain qualitative data for interpretation of KABP data
- separate post-coital interviews with both participants
- decreases in STD incidence among those who report consistent condom use
- inspection of a vaginal smear for evidence of sperm
- count or inspection of used condoms.

The issue of bias in survey data will be addressed in more detail in a future tools module.



Use of multiple evaluation methods

Qualitative and quantitative complementary methods

In order to obtain the highest quality results, evaluation activities at all stages should incorporate both qualitative and quantitative methodologies.

- Qualitative data concentrate on words and non-structured observations to express reality, resulting in in-depth information of high validity about fewer cases (ie. respondents, interviewees, people)
- Qualitative research involves small numbers of non-randomly sampled respondents, and does not attempt to generalize results to a larger population
- Qualitative information provides insight into the attitudes, beliefs, motives and behaviors of a target population, with depth of understanding about what people think and how they feel
- It attempts to answer “why” questions and deals with the emotional and contextual aspects of response, adding “feel”, “texture” and nuance to quantitative findings

In qualitative investigation, sample selection often occurs through a process of “network”, “snowball”, or “convenience” sampling where the investigator follows leads and referrals from one respondent to the next resulting in a sampling strategy that is open-ended and opportunistic. Still, care must be taken to ensure that respondents are typical of the various sections of the populations that need to be studied.

Qualitative research can:

- serve as a tool for generating ideas
Example: Focus group discussions with key groups of people around the question of ways to accurately elicit information on actual condom use in a particular subpopulation. Key groups might include researchers at a local research institution, groups of community health workers, or a peer counseling team from a target group.

- be a preliminary step in the development of a quantitative study
 - Example:** In order to determine how to list categories of union between two people in a baseline country-wide survey - besides the usual “marital status” category - several key informants might be asked to list indigenous language names for terms like husband, wife, partner, single person, married, widowed, divorced, separated, visiting, girlfriend/boyfriend, and other potential descriptive terms for ways that two people might be in union.
- help explain the results of a quantitative study

Suppose within a particular target area, condom distribution statistics show a steady increase, but STD rates remain the same or increase. Qualitative methods such as participant observation, key informant interviewing, or focus group discussions might be used to further investigate the apparent conflict between sets of quantitative data.

- be a primary data collection method
 - Example:** In order to investigate health seeking behavior for STDs to determine the factors influencing peoples’ decisions to consult various traditional or biomedical resources, qualitative methods would yield more accurate data. Information about use of traditional or alternative therapeutic care is a sensitive topic which people often will not discuss with an interviewer they don’t know personally. People often use traditional, alternative and biomedical resources concurrently or sequentially, but in a survey format will only admit to use of biomedical facilities.)

In contrast, quantitative data place considerable trust in numbers that represent opinions or concepts.

- They produce more breadth of information across a larger number of cases, and are characterized by high reliability and replicability
- Quantitative approaches provide a *measurement* of informant response, answering questions of “how many” or “how often”, and are based on statistically appropriate samples of target populations
- Quantitative data collection can be standardized between communities, countries and time periods, and the resulting sampled information can be generalized to larger populations, depending on the sampling procedures used
- AIDS prevention programs have used quantitative methods such as pre- and post-intervention knowledge, attitudes, beliefs and practices (KABP) surveys, surveillance data, and case-control studies for developing and evaluating interventions

The key to effectively integrating the two methodological strategies is by carefully selecting sub-samples of respondents or participants for intensive qualitative investigation from a larger quantitative sample population, or by ensuring that all of the regional and social categories being studied with quantitative methods are included in the in-depth/qualitative sample.

While AIDS prevention program evaluation is not primarily an epidemiological investigation, its goal should be to determine the program's effectiveness in reducing the incidence of HIV. To ensure that program managers are not overburdened, program evaluation strategies should attempt to integrate data collection into the daily activities of program staff so that it becomes routine. Evaluation strategies should also be carefully planned to ensure that maximum use is made of project data, and that data use is coordinated among researchers, epidemiologists, managers, and evaluators.

An integrated evaluation strategy is based upon a set of preferred (first choice) and back-up indicators. HIV incidence among the population targeted by the intervention is the preferred indicator of changes in sexual transmission of HIV. If it is not feasible to collect incidence data, national sentinel surveillance data for specific population groups in particular sites can be used to detect changes and monitor trends in the prevalence of HIV infection (Slutkin 1988).

Findings based on populations such as STD patients or antenatal clients can be informative for discerning HIV trends in a target population, as long as there are no major fluctuations in the utilization of the sentinel site by the target population (ie., as long as differences in the demographic profiles of clients compared to the general population are taken into consideration). If sentinel surveillance data are used to monitor trends in HIV infection, the comparability of the intervention program's target population and the sentinel surveillance population should be assessed. Evaluation activities can be linked to a country's sentinel surveillance program by asking those tested about exposure to the AIDS prevention program.

Methods for assessing changes

This section reviews three methods of data collection helpful in assessing changes in condom use, partner acquisition and exposure to the program. These three methods have proven useful in evaluating peer education/condom distribution programs, especially where HIV and STD data are not available and long surveys are inappropriate or not feasible:

- Periodic cross-sectional short quantitative KABP surveys focusing on risk behaviors, condom use and program exposure
- Process indicator data on intervention outputs or activities, ie., numbers of condoms distributed, people trained, and educational materials distributed
- Periodic focus group discussion and in-depth interview data on attitudes, condom acceptability, educational materials and logistics

Periodic cross-sectional short surveys can provide quantitative estimates of a wide variety of indicators, including:

- the prevalence of high risk behaviors such as sex with multiple partners
- the level of condom use and its association with program exposure
- the extent of program coverage and awareness

Questions included in these surveys may be divided into the following categories of variables:

- socio-demographic information
- knowledge of AIDS/HIV
- protective behaviors (use of condoms and others)
- level of exposure to the program
- risky behaviors (drug use, multiple partners, etc)

Process indicator records maintained by program implementation staff are quantitative indicators used to record program activities, monitor condom distribution, and identify problems. Process indicators, or service statistics, are items such as number of:

- patients seen at the clinic,
- peer educators trained,
- brochures produced and distributed
- educational contacts made by program staff

These statistics provide insight into the question of what services the intervention actually provided.

For example, a chart of condom sales per quarter shows whether the condom social marketing program has found a market for condoms. When used with survey data regarding the source of condoms, condom sales data can provide important evidence of the penetration of the intervention into the target community.

Periodic focus groups and in-depth interviews provide qualitative data which can be used to interpret quantitative data, develop educational materials, and identify barriers to condom use. The success of an intervention may be related to variables that are difficult to measure in surveys, such as the quality of interaction between health care providers and individuals practicing high-risk behaviors, or the degree of access to and cooperation of the target population.

Effective use of multiple evaluation methods reduces the bias arising from using any one method alone, and permits cross-validation of results.

Examples of the use of multiple methods

Even if it's not be possible to fully integrate methods due to budget or time constraints, it's usually possible to at least compare the information obtained from two methods. This provides validation of the information and gives insight into how well the program is being implemented.

Examples from Yaounde, Cameroon and Bamako, Mali have used focus group data to help interpret quantitative KABP estimates of condom use. The table below presents Cameroon sex workers' responses to survey and focus group questions about frequency of condom use with clients. About half of the respondents reported using condoms less than half of the time, or never. Insight into the initial low level of condom use was provided by focus group data indicating that there were social, economic, and practical barriers to the use of condoms.

Using KABP and Focus Group Data

KABP: How often do you use a condom with your clients?

| | |
|-------------------------|-----------|
| Never | 18 |
| Less than half the time | 31 |
| About half the time | 15 |
| Always | 20 |
| Missing | 5 |
| Total | 89 = 100% |

FOCUS GROUP: "The methods of preventing AIDS are well known but using condoms pose numerous problems ... available only in pharmacies ... excessive cost ... discomfort ... client refusal."

The next table shows how process data can be used to validate results from KABP data. The number of condoms used by 82 program participants is estimated based on their answers to a KABP question about frequency of condom use. The estimate of 62,335 condoms is then compared with process information that indicates 58,000 condoms were distributed by the program. Therefore, based on the agreement between the two measures, we conclude that the reported rates of condom use are probably accurate.

Using KABP and Process Data

KABP Data: Condom use among 82 program participants

| | | |
|-------------|----------------------------------|-----------------|
| "Always" | (55 women x 5 acts x 182 days) | = 50,050 |
| "Sometimes" | (27 women x 2.5 acts x 182 days) | = 12,285 |
| | ESTLMATED TOTAL CONDOMS | = 62,335 |

Process Data: condoms distributed over 6 months = 58,000

In conclusion

Future modules in this series will feature techniques for conducting focus group discussions, guidelines for doing rapid ethnographic studies, overviews of quantitative methods for evaluation research, monitoring projects using Process Indicator Forms, evaluating capacity building or institutional strengthening activities, and other topics as the need arises. AIDSCAP regional evaluation officers and headquarters evaluation staff can provide clarification of issues raised in this and other modules. The AIDSCAP headquarters evaluation unit would also like to receive feedback from readers, especially HIV/AIDS prevention program managers, on ways to improve future editions of this introductory module, or any of the other modules in the series.

SECTION



10

Appendices

References and bibliography

- Anderson, R.M. and R.M. May
1988 Epidemiological parameters of HIV transmission, *Nature* 133:514-519.
- Bentley, M.E., G.H. Pelto, W.L. Straus et al.
1988 Rapid ethnographic assessment: applications in a diarrhea management program, *Social Science and Medicine* 27(1):107-16.
- Cantania, J.S., D.R. Gibson, D.D. Chitwood, T.J. Coates
1990 Methodological problems in AIDS behavioral research: influences on measurement error and participation bias in studies of sexual behavior, *Psychological Bulletin* 108(3):339-362.
- Coreil, J., A. Augustin, E. Holt and N.A. Halsey.
1989 Use of ethnographic research for instrument development in a case-control study of immunization use in Haiti, *International Journal of Epidemiology* 19 (4 suppl 2):S33-7.
- Coyle, S., R.F. Boruch, C.F. Turner, editors
1991 *Evaluating AIDS Prevention Programs: Expanded Edition*. National Research Council. Washington, DC: National Academy Press.
- Debus, M.
1986 *Handbook for Excellence in Focus Group Research*. Academy for Educational Development/HealthCom, 1255 23rd St, NW, Washington, DC 20037 and Porter/Novelli 1001 30th St, NW, Washington, DC 20007.

- Fisher, A.A., J.E. Laing, J.E. Stoeckel, and J.W. Townsend
 1991 Handbook for Family Planning Operations Research Design. Second Edition. Population Council, New York.
- Holmes K.K., P.Mardh, P.F. Sparling, P.J. Wiesner, W. Cates, S.W. Lemon, W.E. Stamm (eds.)
 1990 Sexually Transmitted Diseases. Second Edition. New York:McGraw Hill.
- James, N.J., C.J. Bignell, P. Gillies
 1991 The reliability of self-reported sexual behavior, *AIDS* 5:333-336.
- Kelsey, J.L., W.D. Thompson, and A.S. Evans
 1986 Methods in Observational Epidemiology. New York: Oxford University Press.
- Kleinbaum, D.G., L.L. Kupper, and H. Morgenstern
 1982 Epidemiologic Research: Principles and Quantitative Methods. Belmont: Lifetime Learning Publications.
- Krueger, R.A.
 1988 Focus Groups: A Practical Guide for Applied Research. Newbury Park: Sage Publications.
- Lamptey, P. and M. Potts
 1990 Targeting of prevention programs in Africa. In P. Lamptey and P. Piot (eds.) The Handbook for AIDS Prevention in Africa. FHI:Durham, pp. 140-180.
- Last, J.M.
 1983 A Dictionary of Epidemiology. Edited for the International Epidemiological Association. Oxford University Press.
- Morgan, D.L.
 1988 Focus Groups as Qualitative Research. Qualitative Research Methods Series Volume 16. Newbury Park: Sage Publications.
- Ngugi, E.M. et al
 1989 Prevention of transmission of human immunodeficiency virus in Africa: Effectiveness of condom promotion and health education among prostitutes, *Lancet* 2:887-890.
- O'Connor, P.
 ND Guidelines on Planning and Conducting Rapid Ethnographic Research: Issues Program Managers Need to Consider. FHI:Durham.

- Patton, M.Q.
 1982 Practical Evaluation. Sage Publications.
 1990 Qualitative Evaluation and Research Methods, Second Edition. Sage Publications.
- Piot, P., M. Caraël
 1988 Epidemiological and sociological aspects of HIV infection in developing countries, *Br Med Bull* 44(1):68-88.
- Schensul, J.J. and S.L. Schensul.
 1990 Ethnographic Evaluation of AIDS Prevention Programs, in *New Directions for Program Evaluation* no. 46 summer.
- Scrimshaw, S.C.M. et al
 1991 The AIDS Rapid Anthropological Assessment Procedures: a Tool for Health Education Planning and Evaluation, *Health Education Quarterly* 18(1):111-123.
 1992 HIV/AIDS Rapid Assessment Procedures: Rapid Anthropological Approaches for Studying AIDS Related Beliefs, Attitudes and Behaviors. United Nations University Food, Nutrition and Development Programme (Harvard Center for Population Studies) 22 Plympton Street, Cambridge, MA 02138.
- Slutkin, G.
 1988 Sentinel surveillance for HIV infection: A method to monitor HIV infection trends in population groups. WHO.
- Stewart, D.W. and P.N. Shamdasani
 1990 Focus Groups: Theory and Practice. Applied Social Research Methods Series Volume 20. Newbury Park: Sage Publications.
- Upchurch, D.M., C. Weisman, M. Shepherd, R. Brookmeyer, R. Fox, D.D. Celentano, L. Colletta, E. Hook
 1991 Interpartner reliability of reporting of recent sexual behavior, *Am J Epidemiol.* 134(10):1159-1166.
- Wilson, D. and A. Mahryar
 1991 The role of AIDS knowledge, attitudes, belief and practice research in sub-Saharan Africa, *AIDS* 5 (suppl 1):S177-181.

Priority Prevention Indicators (PPIs)

WHO/Global Program on AIDS (GPA)
(as of 7 May 1993)

PPI 1: KNOWLEDGE OF PREVENTIVE PRACTICES

Number of people citing at least two
acceptable ways of protection from HIV infection
Population aged 15 - 49 reporting

PPI 2: CONDOM AVAILABILITY (Central level)

Total number of condoms available
for distribution during the preceding 12 months
Population aged 15 - 49

PPI 3: CONDOM AVAILABILITY (Peripheral level)

Number of people who can acquire a condom
Population aged 15 - 49

PPI 4: REPORTED NON-REGULAR SEXUAL PARTNERS

Number of people aged 15 - 49 who report having had at least one
sex partner other than their regular sex partner(s) in the last 12 months
Total number of people aged 15 - 49 who report having been sexually active
in the last 12 months

**PPI 5: REPORTED CONDOM USE IN THE MOST RECENT
SEXUAL INTERCOURSE OF RISK**

Number of people aged 15 - 49 reporting
the use of a condom during the most recent act of sexual intercourse of risk
Total number of people aged 15 - 49 who report sexual intercourse of risk
in the last 12 months

PPI 6: STD CASE MANAGEMENT

Number of individuals presenting with STD in health facilities assessed and
treated in an appropriate way (according to national standards)
Total number of individuals presenting with STD in health facilities

PPI 7: STD CASE MANAGEMENT

Number of individuals seeking STD care in health facilities
who received appropriate advice on condoms and on partner notification
Total number of individuals seeking STD care in health facilities

- PPI 8: STD PREVALENCE, WOMEN (under development)
- $$\frac{\text{Number of pregnant women aged 15 - 24 with positive serology for syphilis}}{\text{Total number of pregnant women aged 15 - 24 attending antenatal clinics whose blood has been screened}}$$
- PPI 9: STD INCIDENCE, MEN
- $$\frac{\text{Number of reported episodes of urethritis in men aged 15 - 49 in the last 12 months}}{\text{Total number of men aged 15 - 49 surveyed}}$$
- PPI 10: HIV PREVALENCE, WOMEN (under development)
- $$\frac{\text{Number of pregnant women aged 15 - 24 seropositive for HIV}}{\text{Total number of pregnant women aged 15 - 24 attending antenatal clinics whose blood has been screened}}$$

