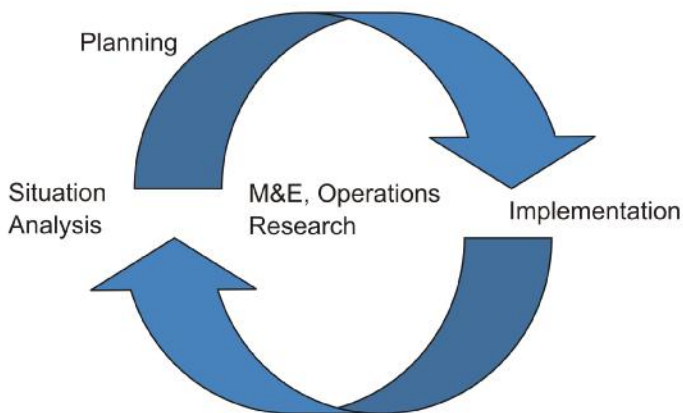


Programme Manager's Planning Monitoring & Evaluation Toolkit



This tool, which was first published in 2004, is subject to constant improvement. We welcome any comments and suggestions you may have on its content. We also encourage you to send us information on experiences from UNFPA-funded and other population programmes and projects that illustrate the issues addressed by this tool.

Please send your inputs to:

**United Nations Population Fund
Viet Nam Country Office**

**1st Floor, UN Apartment Building
2E, Van Phuc Compound, Ha Noi**

Telephone: (84.4) 38236632

Fax: (84.4) 38232822

E-mail: unfpa-fo@unfpa.org.vn

The tool is posted on the UNFPA website at <http://vietnam.unfpa.org/>

TABLE OF CONTENTS

Tool Number 1	
Glossary of Planning, Monitoring & Evaluation Terms	1
Tool Number 2	13
Defining Evaluation	15
Tool Number 3	21
Purposes of Evaluation	23
Tool Number 4	29
Stakeholder Participation in Monitoring and Evaluation	31
Tool Number 5	43
Planning and Managing an Evaluation	45
Part I : Planning Evaluations	45
Part II : Defining Evaluation Questions and Measurement Standards	52
Part III : The Data Collection Process	64
Part IV : Managing the Evaluation Process	75
Part V : Communicating and Using Evaluation Results	87
Part VI : Evaluation Standards	90
Tool Number 6	95
Programme Indicators	97
Part I : Identifying Output Indicators The Basic Concepts	97
Part II : Indicators for Reducing Maternal Mortality	109

Tool Number **1**

**GLOSSARY OF PLANNING,
MONITORING &
EVALUATION TERMS**

An abstract 3D wireframe graphic in shades of orange and yellow, featuring a grid pattern and several curved, mesh-like structures that resemble stylized buildings or data visualizations.

March 2004

GLOSSARY OF PLANNING, MONITORING & EVALUATION TERMS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

The glossary responds to the need for a common understanding and usage of results based planning, monitoring and evaluation terms among UNFPA staff and its partners. In this context, the planning, monitoring and evaluation terminology has been updated to incorporate the definition of terms adopted by the UN Task Force on Simplification and Harmonization.

II. The Glossary

(A)

Accountability: Responsibility and answerability for the use of resources, decisions and/or the results of the discharge of authority and official duties, including duties delegated to a subordinate unit or individual. In regard to programme managers, the responsibility to provide evidence to stakeholders that a programme is effective and in conformity with planned results, legal and fiscal requirements. In organizations that promote learning, accountability may also be measured by the extent to which managers use monitoring and evaluation findings.

Achievement: A manifested performance determined by some type of assessment.

Activities: Actions taken or work performed through which inputs such as funds, technical assistance and other types of resources are mobilized to produce specific outputs.

Analysis: The process of systematically applying statistical techniques and logic to interpret, compare, categorize, and summarize data collected in order to draw conclusions.

Appraisal: An assessment, prior to commitment of support, of the relevance, value, feasibility, and potential acceptability of a programme in accordance with established criteria.

Applied Research: A type of research conducted on the basis of the assumption that human and societal problems can be solved with knowledge. Insights gained through the study of gender relations for example, can be used to develop effective strategies with which to overcome, socio-cultural barriers to gender equality and equity. Incorporating the findings of applied research into programme design therefore can strengthen interventions to bring about the desired change.

Assumptions: Hypotheses about conditions that are necessary to ensure that: (1) planned activities will produce expected results; (2) the cause effect relationship between the different levels of programme results will occur as expected. Achieving results depends on whether or not the assumptions made prove to be true. Incorrect assumptions at any stage of the results chain can become an obstacle to achieving the expected results.

Attribution: Causal link of one event with another. The extent to which observed effects can be ascribed to a specific intervention.

Auditing: An independent, objective, systematic process that assesses the adequacy of the internal controls of an organization, the effectiveness of its risk management and governance processes, in order to improve its efficiency and overall performance. It verifies compliance with established rules, regulations, policies and procedures and validates the accuracy of financial reports.

Authority: The power to decide, certify or approve.

(B)

Baseline Information: Facts about the condition or performance of subjects prior to treatment or intervention.

Baseline Study: An analysis describing the situation prior to a development intervention, against which progress can be assessed or comparisons made.

Benchmark: Reference point or standard against which progress or achievements can be assessed. A benchmark refers to the performance that has been achieved in the recent past by other comparable organizations, or what can be reasonably inferred to have been achieved in similar circumstances.

Beneficiaries: Individuals, groups or entities whose situation is supposed to improve (the target group), and others whose situation may improve as a result of the development intervention.

Bias: Refers to statistical bias. Inaccurate representation that produces systematic error in a research finding. Bias may result in overestimating or underestimating certain characteristics of the population. It may result from incomplete information or invalid data collection methods and may be intentional or unintentional.

(C)

Capacity: The knowledge, organization and resources needed to perform a function.

Capacity Development: A process that encompasses the building of technical abilities, behaviours, relationships and values that enable individuals, groups, organizations and societies to enhance their performance and to achieve their development objectives over time. It progresses through several different stages of development so that the types of interventions required to develop capacity at different stages vary. It includes strengthening the processes, systems and rules that

shape collective and individual behaviours and performance in all development endeavours as well as people's ability and willingness to play new developmental roles and to adapt to new demands and situations. Capacity development is also referred to as capacity building or strengthening.

Causality Analysis: A type of analysis used in programme formulation to identify the root causes of development challenges. Development problems often derive from the same root causes (s). The analysis organizes the main data, trends and findings into relationships of cause and effect. It identifies root causes and their linkages as well as the differentiated impact of the selected development challenges. Generally, for reproductive health and population problems, a range of causes can be identified that are interrelated. A "causality framework or causality tree analysis" (sometimes referred to as "problem tree") can be used as a tool to cluster contributing causes and examine the linkages among them and their various determinants.

Chain of Results: The causal sequence in the planning of a development intervention that stipulates the possible pathways for achieving desired results beginning with the activities through which inputs are mobilized to produce specific outputs, and culminating in outcomes, impacts and feedback. The chain of results articulates a particular programme theory.

Conclusion: A reasoned judgement based on a synthesis of empirical findings or factual statements corresponding to a specific circumstance.

Cost-Benefit Analysis: A type of analysis that compares the costs and benefits of programmes. Benefits are translated into monetary terms. In the case of an HIV infection averted, for instance, one would add up all the costs that could be avoided such as medical treatment costs, lost income, funeral costs, etc. The cost-benefit ratio of a programme is then calculated by dividing those total benefits (in monetary terms) by the total programme cost (in monetary terms). If the benefits as expressed in monetary terms are greater than the money spent on the programme, then the programme is considered to be of absolute benefit. Cost-benefit analysis can be used to compare interventions that have different outcomes (family planning and malaria control programmes, for example). Comparisons are also possible across sectors. It is, for instance, possible to compare the cost-benefit ratio of an HIV prevention programme with that of a programme investing in girls' education. However, the valuation of health and social benefits in monetary terms can sometimes be problematic (assigning a value to human life, for example).

Cost-Effectiveness Analysis: A type of analysis that compares effectiveness of different interventions by comparing their costs and outcomes measured in physical units (number of children immunized or the number of deaths averted, for example) rather than in monetary units. Cost-effectiveness is calculated by dividing the total programme cost by the units of outcome achieved by the programme (number of deaths averted or number of HIV infections prevented) and is expressed as cost per death averted or per HIV infection prevented, for example. This type of analysis can only be used for programmes that have the same objectives or outcomes. One might compare, for instance, different strategies to reduce maternal mortality. The programme that costs less per unit of outcome is considered the more cost-effective. Unlike cost-benefit analysis, cost-effectiveness analysis does not measure absolute benefit of a programme. Implicitly, the assumption is that the outcome of an intervention is worth achieving and that the issue is to determine the most cost-effective way to achieve it.

Coverage: The extent to which a programme reaches its intended target population, institution or geographic area.

(D)

Data: Specific quantitative and qualitative information or facts.

Database: An accumulation of information that has been systematically organized for easy access and analysis. Databases are usually computerized.

(E)

Effectiveness: A measure of the extent to which a programme achieves its planned results (outputs, outcomes and goals).

Effective Practices: Practices that have proven successful in particular circumstances. Knowledge about effective practices is used to demonstrate what works and what does not and to accumulate and apply knowledge about how and why they work in different situations and contexts.

Efficiency: A measure of how economically or optimally inputs (financial, human, technical and material resources) are used to produce outputs.

Evaluability: The extent to which an activity or a programme can be evaluated in a reliable and credible fashion.

Evaluation: A time-bound exercise that attempts to assess systematically and objectively the relevance, performance and success, or the lack thereof, of ongoing and completed programmes. Evaluation is undertaken selectively to answer specific questions to guide decision-makers and/or programme managers, and to provide information on whether underlying theories and assumptions used in programme development were valid, what worked and what did not work and why. Evaluation commonly aims to determine the relevance, validity of design, efficiency, effectiveness, impact and sustainability of a programme.

Evaluation Questions: A set of questions developed by the evaluator, sponsor, and/or other stakeholders, which define the issues the evaluation will investigate and are stated in such terms that they can be answered in a way useful to stakeholders.

Evaluation Standards: A set of criteria against which the completeness and quality of evaluation work can be assessed. The standards measure the utility, feasibility, propriety and accuracy of the evaluation. Evaluation standards must be established in consultation with stakeholders prior to the evaluation.

Evaluative Activities: Activities such as situational analysis, baseline surveys, applied research and diagnostic studies. Evaluative activities are quite distinct from evaluation; nevertheless, the findings of such activities can be used to improve, modify or adapt programme design and implementation.

Ex-ante Evaluation: An evaluation that is performed before implementation of a development intervention. Related term: appraisal.

Ex-post Evaluation: A type of summative evaluation of an intervention usually conducted after it has been completed. Its purpose is to understand the factors of success or failure, to assess the outcome, impact and sustainability of results, and to draw conclusions that may inform similar interventions in the future.

Execution: The management of a specific programme which includes accountability for the effective use of resources.

External Evaluation: An evaluation conducted by individuals or entities free of control by those responsible for the design and implementation of the development intervention to be evaluated (Synonym: independent evaluation).

(F)

Feasibility: The coherence and quality of a programme strategy that makes successful implementation likely.

Feedback: The transmission of findings of monitoring and evaluation activities organized and presented in an appropriate form for dissemination to users in order to improve programme management, decision-making and organizational learning. Feedback is generated through monitoring, evaluation and evaluative activities and may include findings, conclusions, recommendations and lessons learned from experience.

Finding: A factual statement on a programme based on empirical evidence gathered through monitoring and evaluation activities.

Focus Group: A group of usually 7-10 people selected to engage in discussions designed for the purpose of sharing insights and observations, obtaining perceptions or opinions, suggesting ideas, or recommending actions on a topic of concern. A focus group discussion is a method of collecting data for monitoring and evaluation purposes.

Formative Evaluation: A type of process evaluation undertaken during programme implementation to furnish information that will guide programme improvement. A formative evaluation focuses on collecting data on programme operations so that needed changes or modifications can be made to the programme in its early stages. Formative evaluations are used to provide feedback to programme managers and other personnel about the programme that are working and those that need to be changed.

(G)

Goal: The higher order objective to which a development intervention is intended to contribute.

(I)

Impact: Positive and negative long term effects on identifiable population groups produced by a development intervention, directly or indirectly, intended or unintended. These effects can be economic, socio-cultural, institutional, environmental, technological or of other types.

Impact Evaluation: A type of outcome evaluation that focuses on the broad, longer-term impact or results of a programme. For example, an impact evaluation could show that a decrease in a community's overall maternal mortality rate was the direct result of a programme designed to improve referral services and provide high quality pre- and post-natal care and deliveries assisted by skilled health care professionals.

Indicator: A quantitative or qualitative measure of programme performance that is used to demonstrate change and which details the extent to which programme results are being or have been achieved. In order for indicators to be useful for monitoring and evaluating programme results, it is important to identify indicators that are direct, objective, practical and adequate and to regularly update them.

Inputs: The financial, human, material, technological and information resource provided by stakeholders (i.e. donors, programme implementers and beneficiaries) that are used to implement a development intervention.

Inspection: A special, on-the-spot investigation of an activity that seeks to resolve particular problems.

Internal Evaluation: Evaluation of a development intervention conducted by a unit and/or individual/s reporting to the donor, partner, or implementing organization for the intervention.

(J)

Joint Evaluation: An evaluation conducted with other UN partners, bilateral donors or international development banks.

(L)

Lessons Learned: Learning from experience that is applicable to a generic situation rather than to a specific circumstance. The identification of lessons learned relies on three key factors: i) the accumulation of past experiences and insights; ii) good data collection instruments; and iii) a context analysis.

Logical Framework (Log Frame): A dynamic planning and management tool that summarizes the results of the logical framework approach process and communicates the key features of a programme design in a single matrix. It can provide the basis for monitoring progress achieved and evaluating programme results. The matrix should be revisited and refined regularly as new information becomes available.

Logical Framework Approach: A specific strategic planning methodology that is used to prepare a programme or development intervention. The methodology entails a participatory process to clarify outcomes, outputs, activities and inputs, their causal relationships, the indicators with which to gauge/measure progress towards results, and the assumptions and risks that may influence success and failure of the intervention. It offers a structured logical approach to setting priorities and building consensus around intended results and activities of a programme together with stakeholders.

(M)

Management Information System: A system, usually consisting of people, procedures, processes and a data bank (often computerized) that routinely gathers quantitative and qualitative information on pre-determined indicators to measure programme progress and impact. It also informs decision-making for effective programme implementation.

Means of Verification (MOV): The specific sources from which the status of each of the results indicators in the Results and Resources Framework can be ascertained.

Meta-evaluation: A type of evaluation that aggregates findings from a series of evaluations. Also an evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators.

Methodology: A description of how something will be done. A set of analytical methods, procedures and techniques used to collect and analyse information appropriate for evaluation of the particular programme, component or activity.

Monitoring: A continuous management function that aims primarily at providing programme managers and key stakeholders with regular feedback and early indications of progress or lack thereof in the achievement of intended results. Monitoring tracks the actual performance against what was planned or expected according to pre-determined standards. It generally involves collecting and analysing data on programme processes and results and recommending corrective measures.

Multi-Year Funding Framework (MYFF): A four-year framework that is composed of three interlinking components: (1) a results framework, which identifies the major results that UNFPA aims to achieve, its key programme strategies, and the indicators that will be used to measure progress; (2) an integrated resources framework that indicates the level of resources required to achieve the stated results; and (3) a managing for results component that defines the priorities for improving the Fund's organizational effectiveness.

Note: For the period 2008-2011, UNFPA developed the Strategic Plan (SP) to serve as the centrepiece for organizational programming, management and accountability.

(O)

Objective: A generic term usually used to express an outcome or goal representing the desired result that a programme seeks to achieve.

Operations Research: The application of disciplined investigation to problem-solving. Operations research analyses a problem, identifies and then tests solutions.

Outcome: The intended or achieved short and medium-term effects of an intervention's outputs, usually requiring the collective effort of partners. Outcomes represent changes in development conditions which occur between the completion of outputs and the achievement of impact.

Outcome Evaluation: An in-depth examination of a related set of programmes, components and strategies intended to achieve a specific outcome. An outcome evaluation gauges the extent of success in achieving the outcome; assesses the underlying reasons for achievement or non achievement; validates the contributions of a specific organization to the outcome; and identifies key lessons learned and recommendations to improve performance.

Outputs: The products and services which result from the completion of activities within a development intervention.

(P)

Participatory Approach: A broad term for the involvement of primary and other stakeholders in an undertaking (e.g. programme planning, design, implementation, monitoring and evaluation).

Performance: The degree to which a development intervention or a development partner operates according to specific criteria/standards/guidelines or achieves results in accordance with stated plans.

Performance Indicator: A quantitative or qualitative variable that allows the verification of changes produced by a development intervention relative to what was planned.

Performance Measurement: A system for assessing the performance of development interventions, partnerships or policy reforms relative to what was planned in terms of the achievement of outputs and outcomes. Performance measurement relies upon the collection, analysis, interpretation and reporting of data for performance indicators.

Performance Monitoring: A continuous process of collecting and analysing data for performance indicators, to compare how well development interventions, partnerships or policy reforms are being implemented against expected results.

Process Evaluation: A type of evaluation that examines the extent to which a programme is operating as intended by assessing ongoing programme operations. A process evaluation helps programme managers identify what changes are needed in design, strategies and operations to improve performance.

Programme: A time-bound intervention similar to a project but which cuts across sectors, themes or geographic areas, uses a multi-disciplinary approach, involves multiple institutions, and may be supported by several different funding sources.

Programme Approach: A process which allows governments, donors and other stakeholders to articulate priorities for development assistance through a coherent framework within which components are interlinked and aimed towards achieving the same goals. It permits all donors, under government leadership, to effectively contribute to the realization of national development objectives.

Programme Theory: An approach for planning and evaluating development interventions. It entails systematic and cumulative study of the links between activities, outputs, outcomes, impact and contexts of interventions. It specifies upfront how activities will lead to outputs, outcomes and longer-term impact and identifies the contextual conditions that may affect the achievement of results.

Project: A time-bound intervention that consists of a set of planned, interrelated activities aimed at achieving defined programme outputs.

Proxy Measure or Indicator: A variable used to stand in for one that is difficult to measure directly.

(Q)

Qualitative Evaluation: A type of evaluation that is primarily descriptive and interpretative, and may or may not lend itself to quantification.

Quantitative Evaluation: A type of evaluation involving the use of numerical measurement and data analysis based on statistical methods.

(R)

Reach: the coverage (e.g., the range or number of individuals, groups, institutions, geographic areas; etc.) that will be affected by a programme.

Recommendation: Proposal for action to be taken in a specific circumstance, including the parties responsible for that action.

Relevance: The degree to which the outputs, outcomes or goals of a programme remain valid and pertinent as originally planned or as subsequently modified owing to changing circumstances within the immediate context and external environment of that programme.

Reliability: Consistency and dependability of data collected through repeated use of a scientific instrument or data collection procedure under the same conditions. Absolute reliability of evaluation data is hard to obtain. However, checklists and training of evaluators can improve both data reliability and validity.

Research: The general field of disciplined investigation.

Result: The output, outcome or impact (intended or unintended, positive and /or negative) derived from a cause and effect relationship set in motion by a development intervention.

Results Based Management (RBM): A management strategy by which an organization ensures that its processes, products and services contribute to the achievement of desired results (outputs, outcomes & impacts). RBM rests on stakeholder participation and on clearly defined accountability for results. It also requires monitoring of progress towards results and reporting on performance/feedback which is carefully reviewed and used to further improve the design or implementation of the programme.

Results Framework: The logic that explains how results are to be achieved, including causal relationships and underlying assumptions. The results framework is the application of the logical framework approach at a strategic level, across an entire organization, for a country programme, a programme component within a country programme, or even a project.

Risks: Factors that may adversely affect delivery of inputs, completion of activities and achievement of results. Many risk factors are outside the control of the parties responsible for managing and implementing a programme.

Risk Analysis: An analysis or assessment of factors that affect or are likely to affect the achievement of results. Risk analysis provides information that can be used to mitigate the impact of identified risks. Some external factors may be beyond the control of programme managers and implementers, but other factors can be addressed with some slight adjustments in the programme strategy. It is recommended that stakeholders take part in the risk analysis as they offer different perspectives and may have pertinent and useful information about the programme context to mitigate the risks.

(S)

Stakeholders: People, groups or entities that have a role and interest in the aims and implementation of a programme. They include the community whose situation the programme seeks to change; field staff who implement activities; and programme managers who oversee implementation; donors and other decision-makers who influence or decide the course of action related to the programme; and supporters, critics and other persons who influence the programme environment (see target group and beneficiaries).

Strategies: Approaches and modalities to deploy human, material and financial resources and implement activities to achieve results.

Success: A favourable programme result that is assessed in terms of effectiveness, impact, sustainability and contribution to capacity development.

Summative Evaluation: A type of outcome and impact evaluation that assesses the overall effectiveness of a programme.

Survey: Systematic collection of information from a defined population, usually by means of interviews or questionnaires administered to a sample of units in the population (e.g. person, youth, adults etc.). Baseline surveys are carried out at the beginning of the programme to describe the situation prior to a development intervention in order to assess progress; Mid line surveys are conducted at the mid point of the cycle to provide management and decision makers with the information necessary to assess and, if necessary, adjust, implementation, procedures, strategies and institutional arrangements, for the attainment of results. In addition, the results of midline surveys can also be used to inform and guide the formulation of a new country programme. End line surveys are conducted towards the end of the cycle to provide decision makers and planners with information with which to review the achievements of the programme and generate lessons to guide the formulation and/or implementation of a new programme/ projects.

Sustainability: Durability of programme results after the termination of the technical cooperation channelled through the programme. Static sustainability - the continuous flow of the same benefits, set in motion by the completed programme, to the same target groups; dynamic sustainability - the use or adaptation of programme results to a different context or changing environment by the original target groups and/or other groups.

(T)

Target Group: The main stakeholders of a programme that are expected to gain from the results of that programme. Sectors of the population that a programme aims to reach in order to address their needs.

Thematic Evaluation: Evaluation of selected aspects or cross-cutting issues in different types of interventions.

Time-Series Analysis: Quasi-experimental designs that rely on relatively long series of repeated measurements of the outcome/output variable taken before, during and after intervention in order to reach conclusions about the effect of the intervention.

Transparency: Carefully describing and sharing information, rationale, assumptions, and procedures as the basis for value judgments and decisions.

(U)

Utility: The value of something to someone or to an institution. The extent to which evaluations are guided by the information needs of their users.

(V)

Validity: The extent to which methodologies and instruments measure what they are supposed to measure. A data collection method is reliable and valid to the extent that it produces the same results repeatedly. Valid evaluations are ones that take into account all relevant factors, given the whole context of the evaluation, and weigh them appropriately in the process of formulating conclusions and recommendations.

(W)

Work Plans: Quarterly, annual, or multiyear schedules of expected outputs, tasks, timeframes and responsibilities.

Sources

OECD/DAC. "**Glossary of Key Terms in Evaluation and Results Based Management**", The DAC Working Party on Aid Evaluation, 2002. Available online at <http://www.oecd.org>

Scriven, Michael. "**Evaluation Thesaurus - Fourth Edition**", Sage Publications, 1991.

The United Nations Development Group (UNDG). "**Results Based Management Terminology**", June 2003. Available online at <http://www.undg.org>

The World Bank. "**Measuring Efficiency and Equity Terms**". Available online at <http://www.worldbank.org>

Tool Number **2**

DEFINING EVALUATION

The background of the lower half of the page is a red-tinted abstract 3D visualization. It features a grid floor and several wireframe structures that resemble peaks or data points, with some elements highlighted in a lighter shade of red. The overall aesthetic is modern and data-driven.

DEFINING EVALUATION

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This tool defines the concept of evaluation, what it is and why we evaluate, the role of evaluation in relation to monitoring and audit, and its role in the context of results-based management approaches (RBM). The content is based on a review of a wide range of evaluation literature from academia and international development agencies such as UNDP, UNICEF, WFP, OECD and bilateral donor agencies such as USAID.

II. What is Programme Evaluation?

Programme evaluation is a **management tool**. It is a time-bound exercise that attempts to **assess systematically and objectively the relevance, performance and success of ongoing and completed programmes and projects**. Evaluation is undertaken selectively to answer specific questions to guide decision-makers and/or programme managers, and to provide information on whether underlying theories and assumptions used in programme development were valid, what worked and what did not work and why. Evaluation **commonly aims to determine the relevance, efficiency, effectiveness, impact and sustainability of a programme or project**².

III. Why evaluate?

The **main objectives** of programme evaluation are:

- To **inform decisions on operations, policy, or strategy** related to ongoing or future programme interventions;
- To **demonstrate accountability**³ to decision-makers (donors and programme countries).

It is expected that improved decision-making and accountability will lead to better results and more efficient use of resources.

¹ This tool was first published in November 2000.

² Definitions of these terms are provided in Tool Number 1: Glossary of Planning Monitoring and Evaluation Terms and are further discussed in Tool Number 5, Part II: Defining Evaluation Questions and Measurement Standards.

³ Accountability is the responsibility to justify expenditures, decisions or the results of the discharge of authority and official duties, including duties delegated to a subordinate unit or individual. Programme Managers are responsible for providing evidence to stakeholders and sponsors that a programme is effective and in conformity with its coverage, service, legal and fiscal requirements.

Other objectives of programme evaluation include:

- To enable *corporate learning* and contribute to the body of knowledge on what works and what does not work and why;
- To *verify/improve* programme *quality* and *management*;
- To *identify successful strategies* for extension/expansion/replication;
- To *modify unsuccessful strategies*;
- To *measure effects/benefits* of programme and project interventions;
- To *give stakeholders the opportunity to have a say* in programme output and quality;
- To *justify/validate* programmes to donors, partners and other constituencies.

IV. What is the Relationship between Monitoring and Evaluation?

Monitoring and evaluation are intimately related. Both are necessary management tools to inform decision-making and demonstrate accountability. Evaluation is not a substitute for monitoring nor is monitoring a substitute for evaluation. Both use the same steps (see **Box 1**), however, they produce different kinds of information. Systematically generated monitoring data is essential for successful evaluations.

Box 1. Evaluation Steps

The evaluation process normally includes the following steps:

- **Defining standards** against which programmes are to be evaluated. In the UNFPA logframe matrix, such standards are defined by the programme indicators;
- **Investigating the performance** of the selected activities/processes/products to be evaluated based on these standards. This is done by an analysis of selected qualitative or quantitative indicators and the programme context;
- **Synthesizing the results** of this analysis;
- **Formulating recommendations** based on the analysis of findings;
- **Feeding recommendations and lessons learned back** into programme and other decision-making processes.

Monitoring continuously tracks performance against what was planned by collecting and analysing data on the indicators established for monitoring and evaluation purposes. It provides continuous information on whether progress is being made toward achieving results (outputs, outcomes, goals) through record keeping and regular reporting systems. Monitoring looks at both pro-

programme processes⁴ and changes in conditions of target groups and institutions brought about by programme activities. It also identifies strengths and weaknesses in a programme. The performance information generated from monitoring enhances learning from experience and improves decision-making. Management and programme implementers typically conduct monitoring.

Evaluation is a periodic, in-depth analysis of programme performance. It relies on data generated through monitoring activities as well as information obtained from other sources (e.g., studies, research, in-depth interviews, focus group discussions, surveys etc.). Evaluations are often (but not always) conducted with the assistance of external evaluators.

Table 1. Characteristics of Monitoring and Evaluation

Monitoring	Evaluation
Continuous	Periodic: at important milestones such as the mid-term of programme implementation; at the end or a substantial period after programme conclusion
Keeps track; oversight; analyses and documents progress	In-depth analysis; Compares planned with actual achievements
Focuses on inputs, activities, outputs, implementation processes, continued relevance, likely results at outcome level	Focuses on outputs in relation to inputs; results in relation to cost; processes used to achieve results; overall relevance; impact; and sustainability
Answers what activities were implemented and results achieved	Answers why and how results were achieved. Contributes to building theories and models for change
Alerts managers to problems and provides options for corrective actions	Provides managers with strategy and policy options
Self-assessment by programme managers, supervisors, community stakeholders, and donors	Internal and/or external analysis by programme managers, supervisors, community stakeholders, donors, and/or external evaluators

Sources: UNICEF, 1991. WFP, May 2000.

⁴ Transformation of inputs into outputs through activities.

V. When do we need Monitoring and Evaluation results during the Programme Cycle?

- During situation analysis and identification of overall programme focus, lessons learned from past programme implementation are studied and taken into account in the programme strategies;
- During programme design, data on indicators produced during the previous programme cycle serve as baseline data for the new programme cycle. Indicator data also enable programme designers to establish clear programme targets which can be monitored and evaluated;
- During programme implementation, monitoring and evaluation ensures continuous tracking of programme progress and adjustment of programme strategies to achieve better results;
- At programme completion, in-depth evaluation of programme effectiveness, impact and sustainability ensures that lessons on good strategies and practices are available for designing the next programme cycle.

VI. What is the relationship between evaluation and audit?

Like evaluation, audit assesses the effectiveness, efficiency and economy of both programme and financial management and recommends improvement. However, the objective and focus of audit differ from that of evaluation.

Box 2. The differing focus of Audit and Evaluation

Evaluation	=	Accountability + Learning
Audit	=	Accountability

Source: UNDP, 1997.

Unlike evaluation, audit does not establish the relevance or determine the likely impact or sustainability of programme results. Audit verifies compliance with established rules, regulations, procedures or mandates of the organization and assesses the adequacy of internal controls. It also assesses the accuracy and fairness of financial transactions and reports. Management audits assess the managerial aspects of a unit's operations.

Notwithstanding this difference in focus, audit and evaluation are both instruments through which management can obtain a critical assessment of the operations of the organization as a basis for instituting improvements.

VII. What is the role of evaluation in RBM?

International development organizations such as UNFPA currently place strong emphasis on national capacity development, good governance and public sector transparency. In this context,

evaluation, together with continuous monitoring of programme and project progress, is an important tool for result-based management. In assessing what works, what does not work and why, evaluation provides information that strengthens organizational decision-making and promotes a culture of accountability among programme implementers. The lessons highlighted through evaluation enable UNFPA to improve programme and organizational performance. Demonstration of more and higher quality results through improved performance can lead to increased funding of UNFPA assisted projects and programmes.

Box 3 outlines, in no particular order of priority, some characteristics and expected benefits of introducing results-based monitoring and evaluation in the Fund.

Box 3. The Expected Benefits of Strengthening Results-based Monitoring and Evaluation in UNFPA

IF

- Senior management is strongly committed to the use of M&E results in decision-making - commitment influences the management style;
- Staff undertake M&E activities and use M&E data at all stages of the programme cycle;
- Staff apply M&E approaches to all areas of UNFPA operations for example in programme, finance, and human resources management;
- Staff engaged in monitoring and evaluation activities strive to pursue objectivity. They make clear the criteria and values on which their judgments are based;
- Staff are held accountable for results and take risks to achieve them;
- Staff apply lessons learned to programme management;
- Staff is recognized by the organization for achieving good results and for their efforts to counteract risks.



THEN

- UNFPA becomes more efficient and better equipped to adapt to a rapidly changing external environment;
- The quality and effectiveness of UNFPA's assistance increases;
- UNFPA and its partners achieve results;
- UNFPA's credibility improves;
- Funding for UNFPA assistance is likely to increase;
- Staff has a heightened sense of achievement and professional satisfaction; productivity improves.

Source: Adapted from UNICEF, 1998.

Sources

OECD. **"Improving Evaluation Practices: Best Practice Guidelines for Evaluation and Background Paper"**, 1999.

Patton, Michael Quinn. **"Utilization- Focused Evaluation - The New Century Text"**, 3rd Edition, Sage Publications, 1997.

Scriven, Michael. **"Evaluation Thesaurus, Fourth Edition"**, Sage Publications, 1991.

UNDP. **"Results-oriented Monitoring and Evaluation - A Handbook for Programme Managers"**, OESP, 1997.

UNFPA. **"Institutionalization of Results-based Management at UNFPA"**, June 2000. Available at www.unfpa.org/results/index.htm

UNFPA. **"RBM at UNFPA"**, ORM, April 2000. Available at www.unfpa.org/results/index.htm

UNICEF. **"Democratic Evaluation: A proposal for strengthening the evaluation function in International Development Organizations"**, Working Document Number 3, May 1998.

UNICEF. **"EVALUATION - A UNICEF Guide for Monitoring and Evaluation - Making a Difference?"**, Evaluation Office, 1991.

USAID. **"Managing for Results at USAID"**, presentation prepared by Annette Binnendijk for the Workshop on Performance Management and Evaluation, New York, 5-7 October, 1998.

WFP. **"WFP Principles and Methods of Monitoring and Evaluation"**, Executive Board Annual Session, Rome, 22-26 May, 2000.

3

Tool Number

PURPOSES OF EVALUATION



PURPOSES OF EVALUATION

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This tool provides an overview of the most frequent reasons for undertaking programme evaluations. The content is based on a review of evaluation literature from academia and international development agencies such as UNFPA, UNDP and UNICEF.

II. Why define the evaluation purpose?

Before evaluating a programme, the reasons for the evaluation should be clearly defined. If the purpose is not clear, there is a risk that the evaluation will focus on the wrong concerns, draw the wrong conclusions and provide recommendations which will not be useful for the intended users of evaluation results.

Experience has shown that when the evaluation manager determines the main purpose of the evaluation together with the intended users of evaluation findings, the chance that the findings will be used for decision-making is greatly increased.

When planning for an evaluation, the evaluation manager should therefore always ask the following questions: Who wants the evaluation? Why do they want it? How do they intend to use it?

III. Three common evaluation purposes

Box 1 highlights the three most common evaluation purposes and a sample of evaluation questions typically asked by the intended users.

¹ This tool was first published in November 2000.

Box 1. Three Common Evaluation Purposes

To improve the design and performance of an ongoing programme - A formative evaluation.

- What are the programme's strengths and weaknesses? What kinds of implementation problems have emerged and how are they being addressed?
- What is the progress towards achieving the desired outputs and outcomes? Are the activities planned sufficient (in quantity and quality) to achieve the outputs?
- Are the selected indicators pertinent and specific enough to measure the outputs? Do they need to be revised? Has it been feasible to collect data on selected indicators? Have the indicators been used for monitoring?
- Why are some implementers not implementing activities as well as others?
- What is happening that was not expected?
- How are staff and clients interacting? What are implementers' and target groups' perception of the programme? What do they like? Dislike? Want to change?
- How are funds being used compared to initial expectations? Where can efficiencies be realized?
- How is the external environment affecting internal operations of the programme? Are the originally identified assumptions still valid? Does the programme include strategies to reduce the impact of identified risks?
- What new ideas are emerging that can be tried out and tested?

To make an overall judgment about the effectiveness of a completed programme, often to ensure accountability - A summative evaluation.

- Did the programme work? Did it contribute towards the stated goals and outcomes? Were the desired outputs achieved?
- Was implementation in compliance with funding mandates? Were funds used appropriately for the intended purposes?
- Should the programme be continued or terminated? Expanded? Replicated?

To generate knowledge about good practices.

- What is the assumed logic through which it is expected that inputs and activities will produce outputs, which will result in outcomes, which will ultimately change the status of the target population or situation (also called the programme theory)?
- What types of interventions are successful under what conditions?
- How can outputs/outcomes best be measured?
- What lessons were learned?
- What policy options are available as a result of programme activities?

IV. Who uses what kind of evaluation findings?

Certain evaluation findings are particularly suited for decision-making by specific users. For example, programme managers and staff of implementing partners need evaluation findings related to the delivery process and progress towards achieving aims. This type of information will help them choose more effective implementation strategies.

Decision-makers who oversee programmes such as policy makers, senior managers and donors, require evaluation findings related to effectiveness, impact and sustainability. This type of information will enable them to decide whether to continue, modify, or cancel the programme or projects.

Data generated through evaluations, which highlight good practices and lessons learned is essential for those engaged in overall policy and programme design.

It is essential to note that *one type of evaluation findings usually constitutes an essential input to produce other types of findings*. For instance, data on programme implementation processes gathered through a formative evaluation, or through monitoring and review activities, is a necessary input to enable analysis of programme impact and to generate knowledge of good practices. When no impact of activities is found, process data can indicate if this occurred because of implementation failure (i.e. services were not provided hence the expected benefit could not have occurred) or theory failure (i.e. the programme was implemented as intended but failed to produce the expected results). Data on implementation processes enable an analysis of which approaches work or do not work and under what conditions. **Box 2** highlights an example of theory failure which has affected the impact of UNFPA's interventions to reduce maternal mortality rates.

Box 2. Programme Theory for Reducing Maternal Mortality

A Thematic Evaluation conducted in 1997-1998 of 7 Safe Motherhood projects supported by UNFPA illustrates that the assumptions or programme theories underlying the strategies adopted were insufficient to achieve project objectives. All of the projects promoted antenatal care (ANC), and four of the projects included training programmes for TBAs. The underlying programme theory was thus that ANC and TBA training are essential strategies to reduce maternal mortality. However, research evidence shows that antenatal care to detect pregnancy-related complications and training of TBAs without appropriate linkages to the formal health system cannot bring about significant reduction in maternal mortality.

The Evaluation therefore concluded that strategies selected to prevent maternal deaths must be based on the most up-to-date technical information. Several basic premises are now widely known with regard to safe motherhood:

- Every pregnancy faces risks;
- A skilled attendant should be present at every delivery;
- Emergency obstetric care should be accessible; and
- More emphasis is needed on care during birth and immediately afterwards. Post-partum care should include the prevention and early detection of complications in both the mother and the new-born.

Source: UNFPA, 1999.

Box 3 illustrates how important it is that managers of UNFPA funded programmes ensure that different types of evaluation findings are produced during a country programme cycle in order to improve the quality of programme related decisions and enable organizational learning.

Box 3. Evaluation findings produced by UNFPA - the present and future requirements.

During the period 1998/1999, **77%** of evaluations undertaken by UNFPA's country offices were **project evaluations** the purpose of which were **to pass overall judgment on project relevance and performance**. They took place at the completion of project implementation and were usually conducted by independent, mostly national consultants. CST experts also participated in a fair number of evaluations.

The remaining **23%** of project **evaluations aimed at improving project design and performance mid-stream**.

During the same period, the Office of Oversight and Evaluation (OOE) conducted four **Thematic Evaluations and studies** in key strategic areas for the Fund such as Safe Motherhood, UNFPA support to HIV/AIDS -related interventions, Implementing the RH vision: Progress and New Directions for UNFPA; and the Impact of Government Decentralization on UNFPA's programming process. These evaluations **aimed mainly at generating knowledge** to enable the Fund to frame overall policies and strategies, which adequately address the varying local contexts in key programme areas.

As the **results-based approach** to programming becomes well established in UNFPA, **process related data** typically collected through continuous programme monitoring, formative evaluations and operations research², as well as data on **good practices and lessons learned**, generated through analysis of results from many evaluations, **will take on increased importance** for providing the critical answers as to what works, what doesn't and why.

Source: Adapted from DP/FPA/2000/10 of 5 May, 2000: Periodic Report of the Executive Director to the Executive Board on Evaluation.

² Operations Research analyses a problem and identifies and then tests possible solutions. The goal is to arrive at models of programme/project implementation that can be replicated elsewhere.

Sources

Patton, Michael Quinn. **"Utilization - Focused Evaluation - The New Century Text"**, 3rd Edition, Sage Publications, 1997.

Rossi, Peter; Freeman, Howard E.; Lipsey, Mark W. **"Evaluation - A Systematic Approach"**, 6th Edition, SAGE Publications. 1999.

UNDP. **"Results-oriented Monitoring and Evaluation - A Handbook for Programme Managers"**, OESP, 1997.

UNFPA document DP/FPA/2000/10 of 5 May, 2000. **"Periodic Report of the Executive Director to the Executive Board on Evaluation"**. Available online in English at <http://www.unfpa.org/exbrd/>

UNFPA. **"Safe Motherhood"**, Evaluation Report Number 15, 1999.

UNICEF. **"EVALUATION - A UNICEF Guide for Monitoring and Evaluation - Making a Difference?"**, Evaluation Office, 1991.

Tool Number

4

**STAKEHOLDER
PARTICIPATION IN
MONITORING AND
EVALUATION**

August 2004¹

STAKEHOLDER PARTICIPATION IN MONITORING AND EVALUATION

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This tool clarifies the significance and different modalities of stakeholder participation in programme monitoring and evaluation. Its content is based on a review of evaluation literature from academia and international development agencies and NGOs such as the Institute of Development Studies, Sussex, UNFPA, UNDP, UNICEF and Catholic Relief Services .

II. What is participatory monitoring and evaluation?

There is no single definition or approach to participatory M&E leaving the field open for interpretation and experimentation. Most of the documented experiences in participatory M&E are from the area of agricultural, environmental and rural development. Experiences in the health and education fields are less readily available.

However, as highlighted in **Box 1**, the principles guiding the participatory approach to M&E clearly distinguishes it from conventional M&E approaches. Participatory M&E also requires a different mindset, acceptance of a different way of conducting M&E.

¹ This tool was first published in March 2001.

² An excellent review of literature on participatory M&E is provided in Estrella 1997.

Box 1. Principles which Distinguish Conventional M&E from Participatory M&E**Conventional M&E:**

- aims at making a judgment on the programme for accountability purposes rather than empowering programme stakeholders
- strives for "scientific" objectivity of M&E findings thereby distancing the external evaluator(s) from stakeholders
- tends to emphasise the needs for information of programme funders and policy makers rather than programme implementers and people affected by the programme
- focuses on measurement of success according to predetermined indicators.

Participatory M&E:

- is a process of individual and collective learning and capacity development through which people become more aware and conscious of their strengths and weaknesses, their wider social realities, and their visions and perspectives of development outcomes. This learning process creates conditions conducive to change and action
- emphasises varying degrees of participation (from low to high) of different types of stakeholders in initiating, defining the parameters for, and conducting M&E
- is a social process of negotiation between people's different needs, expectations and worldviews. It is a highly political process which addresses issues of equity, power and social transformation
- is a flexible process, continuously evolving and adapting to the programme specific circumstances and needs.

Source: Estrella, 1997.

III. Who are the stakeholders?

M&E stakeholders are those people who have a stake in the programme. They are persons who take decisions using the M&E data and findings.

Box 2 shows five types of stakeholders. They can include members of the community - men, women and youth; health clinic staff, teachers of population education, staff of the Census Bureau who implement the programme activities; national counterparts in government and NGOs at the central and local levels who oversee programme implementation; international and national programme funders and other decision-makers; community leaders, central and local government administrators who have a major influence on the "enabling" programme environment.

Box 2. Types of Stakeholders

- The community whose situation the programme seeks to change
- Project Field Staff who implement activities
- Programme Managers who oversee programme implementation
- Funders and other Decision-Makers who decide the course of action related to the programme
- Supporters, critics and other stakeholders who influence the programme environment.

Source: Adapted from C.T. Davies, 1998.

IV. The rationale for stakeholder participation in M&E

The growing interest within the international aid community in participatory approaches to development programming emanates from lessons learned in the past. It was found that participation of the programme stakeholders, central level decision makers, local level implementers, and communities affected by the programme, in programme design, implementation, monitoring and evaluation, improves programme quality and helps address local development needs. It increases the sense of national and local ownership of programme activities and ultimately promotes the likelihood that the programme activities and their impact would be sustainable (see **Box 3**).

Box 3. Advantages of Stakeholder Participation in M&E Planning and Implementation.

- Ensures that the M&E findings are relevant to local conditions;
- Gives stakeholders a sense of ownership over M&E results thus promoting their use to improve decision-making;
- Increases local level capacity in M&E which in turn contributes to self-reliance in overall programme implementation;
- Increases the understanding of stakeholders of their own programme strategy and processes; what works, does not work and why;
- Contributes to improved communication and collaboration between programme actors who are working at different levels of programme implementation;
- Strengthens accountability to donors;
- Promotes a more efficient allocation of resources.

Sources: Aibel, 1999. UNDP, 1997.

The introduction in UNFPA of the results-based approach to programme management calls for strengthening partnerships, participation and teamwork at all levels and stages of the programme process. Therefore, efforts should be made to move away from the conventional to more participatory approaches to M&E.

However, exactly what programme stakeholders are involved in M&E varies according to the purpose of M&E and the general institutional receptiveness to the use of participatory approaches. In each instance, programme managers must decide which group of stakeholders should be involved, to what extent and how.

V. When is it appropriate to use participatory M&E approaches?

In general, all relevant counterparts such as project field staff, programme managers as well as the UNFPA Country Office should regularly monitor programme activities.

The extent of stakeholder participation in evaluation, however, depends on the evaluation questions and circumstances. Participatory evaluations are particularly useful when there are questions about implementation difficulties or programme effects on different stakeholders or when information is wanted on stakeholders' knowledge of programme goals or their view of progress. A conventional approach to evaluation may be more suitable when there is a need for independent outside judgment and when specialized information is needed that only technical experts can provide. Such an approach is also more appropriate when key stakeholders don't have time to participate, or when such serious lack of agreement exists among stakeholders that a collaborative approach is likely to fail.

Participatory M&E is useful for:

- ***institutional learning and capacity development***: through self-assessment, stakeholders identify and solve programme related problems themselves thereby strengthening their capacity to be active participants in programme implementation, rather than remaining passive recipients of development assistance. Self-assessment can help strengthen partnerships between different stakeholders and increases their understanding of programme processes and outcomes. It also clarifies the roles of different stakeholder in implementing the programme. **Box 4** provides a few lessons from Madagascar on the participation of a key stakeholder group, health service providers, in monitoring the quality of service delivery by using the COPE³ approach.

³ Client-oriented, Provider-efficient. A COPE Handbook can be obtained from AVSC International. For more information on COPE, visit <http://www.engenderhealth.org>

- ***negotiating stakeholder perspectives:*** participatory M&E may be used as a process that allows different stakeholders to articulate and present their needs, interests and expectations. The process of dialogue and negotiation among stakeholders used in participatory M&E facilitates reconciliation of different stakeholder viewpoints. Difficulties may, however, arise in resolving competing and conflicting stakeholder perceptions, especially when certain stakeholder groups are powerless in relation to others.
- ***ensuring public accountability:*** participatory M&E can be a way for programme participants and local citizens themselves to monitor and evaluate the performance of donor and government institutions. For instance, legal reforms that decentralize decision-making often encourage elected representatives at district or municipal levels to be more proactive in monitoring implementation of local development plans. In Paraguay, UNFPA is funding a project the aim of which is to establish a network of local male and female leaders who will monitor the quality of RH service delivery and periodically report on the status of the services to higher levels in the health administration.

Box 4. Lessons from applying the COPE methodology in Madagascar

The COPE approach involves a series of activities (self-assessment, client interviews, client-flow analysis, and action planning) conducted by staff at health service delivery points to help them identify and solve their own problems and continually improve service quality, using resources already available at their facility.

In 1994 the NGO SALFA introduced COPE in Madagascar. By 1998 the method was used by 20 provincial level FP centres managed by SALFA and one government health centre. The experience showed that the method can be applied to many service delivery management processes at clinics and hospitals: for example in the areas of administration and service delivery such as management of staff and supplies and provision of preventive and curative services. The opportunity for service providers to contribute to the assessment and improvement of service delivery increased their sense of responsibility and the enthusiasm for their work. The self-assessment process increased their awareness of the importance of better client provider relations. As a result, service utilization improved significantly.

The introduction of COPE is, however, not problem free:

- Health care providers tended to think that COPE would resolve all their problems, including the lack of financial incentives for good service delivery. The introduction of COPE should therefore emphasise its main purpose of achieving client satisfaction;
- In their enthusiasm, health care providers tended to seek the perfect solutions even if sometimes too ambitious. Therefore, slow achievement of results discouraged them and they lost faith in the utility of the COPE approach. It is important to ensure that solutions proposed can be implemented by health care providers themselves, within the health care facility and with reasonable resources;
- Clients interviewed by service providers did not express all their opinions about the services, leaving out negative aspects. This COPE tool should therefore be applied by data collectors independent from the service delivery point while ensuring client confidentiality;
- The short-term results achieved with the introduction of COPE were not maintained at service delivery points that did not continuously monitor the use of the COPE approach. Continuous monitoring of COPE implementation is key to achieving the expected results;
- Health care providers at the government facility were demoralized by the fact that their supervisors rarely visited their health centre, despite official recognition of their excellent work. Continuous supervision of COPE implementation by higher level supervisors is important to sustain and improve results;
- These health care providers also realized that decisions to solve several of the problems identified needed to be taken at higher levels of the health administration. The introduction of COPE at the individual service delivery point should therefore be combined with the necessary related interventions at higher decision-making levels.

Source: "Le Processus 'COPE' - L'Expérience de Madagascar", UNFPA, Madagascar, 2000

VI. Which stakeholders should participate in evaluation and what role should they play?

Participation may be broad to include a wide array of programme staff, communities affected by the programme, partners and others. It may, alternatively, target one or two of these groups. For example, if the aim is to uncover what hinders programme implementation, field implementers may need to be involved. If the issue is the impact of a programme on local communities, they may be the most appropriate participants. If the aim is to know if all stakeholders understand a programme's goals and view progress similarly, broad participation may be best.

Roles may range from serving as a resource or informant to participating fully in some or all phases of the evaluation.

How can communities be involved in evaluation? Community participation can be constrained by lack of literacy skills, insufficient time, the intensity of analytical work to be undertaken during the evaluation, and the fact that many of the issues covered during the evaluation are not directly relevant to community members. There are different ways to make sure that the community perspective is considered. For instance, prior to a programme evaluation, complementary evaluation activities could be undertaken with communities involved in and affected by the programme. Such activities could include interviews with and collection of data by community members. They could also consist of community members using PRA and PLA tools⁴ to analyse programme activities and assess whether they meet their needs. Alternatively, community members could define their own criteria for evaluating community-based activities and use these criteria to carry out their own evaluation.

Table 1 illustrates responsibilities of participants in an "expert-driven", conventional evaluation process as compared to a participatory evaluation process involving programme managers, field staff and other decision-makers. The example recognizes the difficulty in simultaneous participation of community and other stakeholders in the evaluation.

⁴ Participatory Reflection and Action (PRA) and Participatory Learning and Action (PLA) methods aim at: allowing community members to express their perceptions, priorities, problems and plans for the future; allowing community development workers to listen to and dialogue with community members in order to better understand their lives, perceptions, problems, priorities and plans for the future. PRA/PLA tools include: community mapping; health problem ranking; body mapping; role plays and stories and other tools

Table 1. Responsibilities of participants in "Expert-driven" and Participatory Evaluations.

Activity	Who is responsible?	
	"Expert-driven" Evaluation	Participatory Evaluation
Define evaluation purpose	Evaluation Planners (UNFPA Country Office) Funder(s)	Small group of stakeholders (2-5 persons) responsible for evaluation coordination throughout (coordinating group). Must include Programme/project managers
Define evaluation objectives, questions and data collection methods	Evaluation Planners (UNFPA Country Office) External evaluator(s)	All selected evaluation stakeholders (Programme/project managers, field staff, other decision-makers as required)
Collect and analyse data	External evaluator(s)	Small group of stakeholders (10-12 persons) divided into teams with a team leader External evaluator(s)
Manage logistics (budgets; field work; equipment and supplies)	Evaluation planners (UNFPA Country Office)	Programme/project staff member (logistics coordinator)
Summarize field work findings	External evaluator(s)	Field work team leaders and a few of the stakeholders involved in data collection and analysis
Formulate lessons learned	External evaluator(s)	All selected stakeholders External evaluator(s)
Summarize evaluation results (findings and lessons)	External evaluator(s)	External evaluator(s) One member of the coordinating group
Develop action plan for implementing evaluation results (findings, conclusions, lessons, recommendations)	Evaluation planners (UNFPA Country Office) in consultation with national counterparts	Small group composed of: programme/project managers, field staff, external evaluator(s)
Write report	External evaluator(s)	Small group of stakeholders, external evaluator(s)
Distribute and discuss report; follow-up on implementation of the Action Plan	Evaluation planners (UNFPA Country Office)	Coordinating group
Develop spirit of collaboration and sharing; coordinate and facilitate all steps of the evaluation	External evaluator(s) Evaluation planners (UNFPA Country Office)	External evaluator(s) Coordination group Evaluation planners (UNFPA Country Office)

Source: Adapted from Aubel, 1999.

VII. What are the steps in a participatory evaluation process?

Box 5 illustrates suggested steps and lessons learned based on a participatory evaluation of 15 social sector, health and agriculture projects implemented by NGOs in Haiti⁵. The individual project evaluations were undertaken over a period of three weeks by teams of three to five NGO staff trained in participatory evaluation approaches.

Box 5. NGO staff can successfully evaluate their own projects

Step 1: a planning meeting gathered 36 NGO staff and several community representatives to answer the following key questions: Was there commitment to undertake a participatory evaluation?; Why undertake an evaluation and what should be the purpose?; When should the evaluation take place?; What indicators should be used?; What resources and support could be expected from the local NGOs?; Who in terms of profile and skills should be involved in the evaluation?; Where should the training of evaluators take place?

Step 2: a four-day participatory evaluation workshop during which 29 NGO staff learned to become participatory evaluation facilitators using PRA and PLA techniques. The workshop aimed at creating an understanding of the difference between participatory and traditional evaluations; awareness of social dynamics and class differences and how evaluation facilitators' attitudes and behaviors can adversely affect others; collective exploration of the attitudes and personal qualities of facilitators; imparting skills on how to get community members to "map" their community to give an inside perspective; how to verify findings using different data collection methods and sources (data triangulation).

Step 3: visits to 15 projects over a two-week period. Each facilitator team visited a total of two projects which were not their own. They spent three days at the project site. They sought answers to detailed evaluation questions that they had identified at the planning meeting. Questions related to major areas of project impact, relationships with other partners, sustainability, efficiency, project management, the role of women and gender equity.

Step 4: collective reflection and dissemination of findings. Each team was responsible for consolidating their community work into a brief project report. At the end of their visit, some teams debriefed community project stakeholders in order to check the reliability of their findings. Each team was responsible for making a presentation of their findings to the larger group. All programme partners were invited to attend a final presentation organized by the facilitators.

.../

⁵ This is a summary of Francoise Coupal, Marie Simoneau. 1997. "Participatory Evaluation: A Case Study of CCIC Humanitarian Fund Projects in Haiti". Mosaic.net. 1997.

Box 5. (cont'd)**Some Lessons Learned*****About the Process:***

- Participatory evaluations are a viable alternative to more traditional approaches even for projects that are not participatory by design. The evaluation was cost-effective. It did not take longer than a traditional evaluation due to the large number of facilitators used. However, additional time would have been beneficial for additional field visits, more preparatory coordination with stakeholders, and for group reflection on findings and lessons learned;
- While the quality of individual reports did vary, inexperienced facilitators can achieve enough participatory evaluation skills to answer evaluation questions. When selecting participatory evaluation facilitators, it is very important to choose persons who are open and willing to learn new methods;
- The impact of the participatory evaluation was significantly greater than that of a traditional evaluation. NGO facilitators' perception of poor, illiterate people changed dramatically from viewing them as helpless beneficiaries to self-sufficient and creative individuals. Many of them now use participatory evaluation tools in their day-to-day work. There has been a wider recognition of the importance of stakeholder participation throughout the project cycle.

What the facilitators learned:

- It is essential to use the PRA/PLA tools; they need to be adapted to individual circumstances, however;
- The trust of individual community members should be earned before presenting the tools. That can take time. The initial strategy for approaching community members is very important;
- The availability of interviewees of both sexes is important;
- PRA methods can be time consuming for busy community members.

Source: Coupal et al. 1997.

VIII. Elements of successful stakeholder participation in evaluation

It is important to bear in mind that the local environment, the socio-political power structures and socio-cultural norms and values, influence the evaluation process.

In this context the following are a few important elements of a successful process of stakeholder participation in evaluation:

- the **support** of programme management (implementing partners) and other direct **decision-makers** for the participatory evaluation approach;
- the **evaluator is committed to the principle of participation; has sufficient group facilitation and mediation skills** to enable effective dialogue and discussion and to ensure that the experience is both participatory and focused;
- a **realistic assessment of stakeholders capacity and willingness to participate** on a full-time basis (partly depending on availability of time, supervisor's support, as well as professional gain);
- **representation of the most important stakeholder interests** related to the programme being evaluated;
- investigation into the "**real**" **interests** and issues of key stakeholders. The "real interests" often differ from those, which are openly expressed. Box 6 highlights an approach to discover "real interests" of stakeholders;
- established procedures for mediating power imbalances among stakeholders.

Box 6. Discovering the real interests of stakeholders through dialogue

Goal of Dialogue:	Stakeholders come to more complete understanding of each other's positions.
Dialogue Process:	A conversation that helps develop alternative perspectives, critical examination.
Evaluator's Identity:	Proposes alternative perspectives, facilitates conversations and critiques, mediates and coaches.

Source: Ryan et al., 2000.

Sources

Aubel, Judy. **"Participatory Program Evaluation Manual - Involving Program Stakeholders in the Evaluation Process"**, Catholic Relief Services, Child Survival and Technical Support Project, Second Edition, December 1999. Available in English, Spanish and French at: <http://www.childsurvival.com/features/bookmarks/pemmanual.cfm>

Coupal Françoise, Simoneau Marie. **"Participatory Evaluation: A Case Study of CCIC Humanitarian Fund Projects in Haiti"**, Mosaic.net, 1997. Available in English at: <http://www.mosaic-net-intl.ca/home.html>

Davies, C.T. **"An introduction to Advocacy"**, Addis Ababa, January 1998.

Estrella, Marisol and Gaventa, John. **"Who Counts Reality? Participatory Monitoring and Evaluation: a Literature Review"**, IDS Working Paper 70, 1997.

Ryan, Katherine E., DeStefano, Lizanne Eds. **"Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation"**, New Directions for Evaluation, A Publication of the American Evaluation Association, Number 85, spring 2000.

UNDP. **"Who are the Question-makers - A Participatory Evaluation Handbook"**, OESP, 1997. Available in English at <http://www.undp.org/eo/documents/who.htm>

UNICEF. **"EVALUATION - A UNICEF Guide for Monitoring and Evaluation - Making a Difference?"**, Evaluation Office, 1991.

USAID. **"Conducting a Participatory Evaluation"**, Performance Monitoring and Evaluation TIPS 1996, Number 1. Center for Development Information and Evaluation. Available in English at: http://www.dec.org/usaaid_eval/#004

More resources on participatory monitoring and evaluation are available at the ELDIS website: <http://www.eldis.org/>



Tool Number **5**

PLANNING AND MANAGING AN EVALUATION



August 2004¹

PLANNING AND MANAGING AN EVALUATION

PART I: PLANNING EVALUATIONS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

Tool number 5 provides guidance on how to plan and manage evaluations during the country programme cycle. It is divided into six parts:

- *Part I* gives an overview of planning evaluations;
- *Part II* discusses the process of defining evaluation questions;
- *Part III* informs about options for data collection;
- *Part IV* discusses the management of the evaluation process including: the division of labour in managing and implementing the evaluation, the development of terms of reference, and the selection of evaluator(s);
- *Part V* proposes options for reporting and communicating evaluation results;
- *Part VI* describes standards that should be used to assess the quality of evaluation work.

The content of Part I is based on a review of evaluation literature of bilateral and international development agencies such as the Centres for Disease Control, Danida, and UNICEF as well as UNFPA project evaluation information.

II. Planning for Evaluations over the Course of the Programme Cycle

Planning for evaluations should be an integral part of country programme design so that timely evaluation information is available to inform decision-making and ensure that UNFPA is able to demonstrate accountability to its stakeholders. Evaluation results are useful for making adjustments in the ongoing programme, or for purposes of designing a new country programme cycle. Careful planning of evaluations and periodic updating of evaluation plans also facilitates their management and contributes to the quality of evaluation results.

In planning evaluation activities, the country office together with key programme stakeholder should decide on:

- **WHY:** the purpose of the evaluations, including who will use the evaluation findings and how;

¹ This tool was first published in December 2000.

- **WHAT:** the main objectives of the evaluation and the questions it should address;
- **HOW:** the data sources and collection methods to be used in the evaluations;
- **WHO:** will undertake the evaluations: what expertise is required? Which evaluations should project stakeholder undertake (an internal evaluation)? Which should be conducted with the assistance of external consultants (national and/or international) and/or with involvement of CST experts? What should be the extent of stakeholder involvement?
- **WHEN:** the timing of each evaluation so that their results in each case or in combination can be used to take important programme related decisions;
- **RESOURCES:** the budget required to implement the evaluation plan.

Evaluations can cover entire country programmes; programme components; thematic areas such as IEC, gender, capacity building strategies and other management issues within the programme; and innovative or pilot projects. The following are a few important criteria to consider in determining the number and focus of evaluations to be undertaken in the course of the programme cycle:

- **the importance of the programme component/thematic area** in terms of resources allocated to it within the country programme;
- **the duration of UNFPA funding** of the particular thematic area. For instance, if population education has been funded for a period longer than five years, an evaluation of all of the activities within this thematic area is advisable in order to ascertain their impact and sustainability;
- **the strategic importance of the activities**, for instance in the case of a pilot or innovative project being tested for possible replication or for policy formulation;
- **the quality and relevance of information collected** through monitoring activities during programme implementation. Availability of high quality process information and data on programme indicators collected during implementation will greatly facilitate evaluation activities. A data-base of good quality programme process information may eventually eliminate the need for very lengthy and costly evaluation exercises and reduce the number of evaluations required for adequate decision-making;
- **the cost of undertaking evaluations** which should be commensurate with their influence on programme decisions;
- **timing of evaluation exercises** to ensure that evaluation results are available in time for important decisions to be taken;

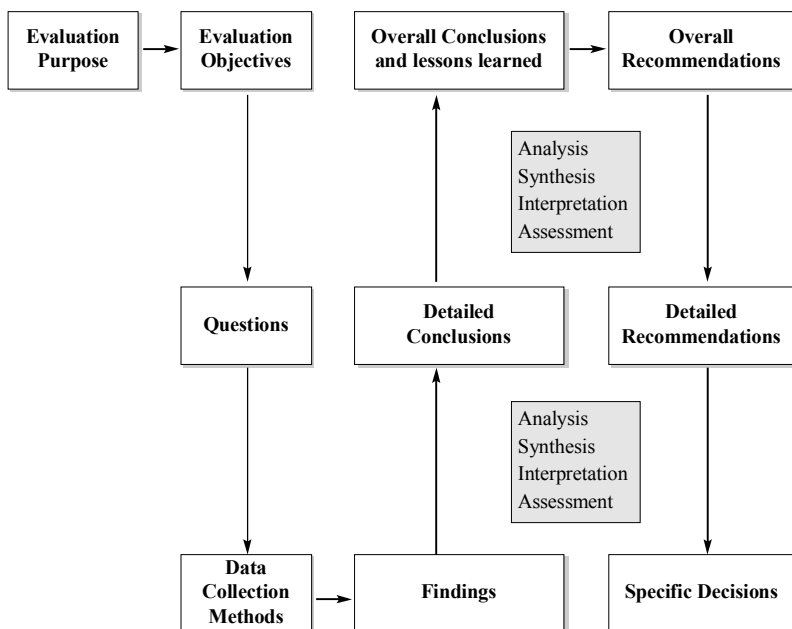
- **the capacity of UNFPA country offices** and programme managers to manage evaluations in light of their workload.

The answer to the question "why undertake an evaluation" is discussed in Tool Number 3 which outlines different purposes for which evaluations could be considered in the course of a programme cycle. The following section gives an overview of the analytical processes through which evaluation objectives and questions ("what") are transformed by using various data collection methodologies ("how") into conclusions and recommendations. Such conclusions and recommendations will allow programme stakeholders, including UNFPA, to take informed country programme management decisions, including those related to funding.

III. The Evaluation Analytical Process

An evaluation analytical process consists of a number of stages as illustrated in Figure 1.

Figure 1. The Evaluation Analytical Process



Source: Adapted from Danida, 1999.

1. Preparation: Defining Evaluation Concerns, Questions and Standards; Selecting Data Collection Methods and Designing Data Collection Instruments.

During the preparatory stage, the evaluation manager and key stakeholders select and define the **evaluation objectives and the questions** to be answered by the evaluation. They also **define standards** to be used to assess the particular programme/thematic area/innovative project being evaluated.

Involving the main users of evaluation results in defining evaluation standards at the beginning of the process is very important as evaluation results will only be used by the intended users if they are confident that the conclusions therein are based on standards that are acceptable to them. Some of these standards, such as the programme results, may have been clarified earlier at the time of programme design or in the course of implementation. However, various stakeholders may interpret details of programme results differently, and results may be redefined over time.

Depending on the nature of the evaluation questions selected, evaluator(s) choose **data collection methods and design instruments** to gather valid and reliable evidence, that enable them to answer the evaluation questions. Data collection instruments include, for example, a list of questions and description of a method for reviewing monitoring information contained in project and programme documents; interview guides for in-depth interviews; focus group guides; and health facility checklists.

Evaluation questions and standards are explained in greater detail in Tool Number 5, part II: Defining Evaluation Questions and Measurement Standards. Data collection methods and instruments are discussed in Tool Number 5, part III: Types of Data and Data Collection Instruments.

2. Implementation: Collecting Information

Information and evidence to answer the evaluation questions are collected using the data collection instruments. The data collected are the evaluation findings.

3. Analysing Evaluation Information; Drawing up Conclusions and Lessons Learned

The evaluation findings are analysed and synthesised into a number of statements, the detailed evaluation conclusions, which provide answers to the evaluation questions. The strength of the conclusions depends on the amount, quality and credibility of the information collected. It also depends on the quality of interpretation and assessment of the evaluation evidence gathered by the evaluator(s).

Interpretation is the effort to determine what the findings mean and is part of the overall effort to make sense of the evidence gathered in an evaluation. Uncovering facts regarding a programme's performance is not sufficient to draw evaluative conclusions. Evaluation evidence must be interpreted to derive the practical significance of what has been learned.

Assessments are statements concerning the merit, worth, or significance of the programme element

being evaluated. They are formed by comparing the findings and interpretations against the selected evaluation standards.

Interpretations and assessments draw on information and perspectives that stakeholders bring to the evaluation inquiry and on the evaluator(s)' expert knowledge on the subject matter.

The detailed evaluation conclusions are then further synthesized into **overall evaluation conclusions and lessons learned**. Lessons are learning from experience. They are generalizations about a specific circumstance. They are formulated drawing both from evidence of the particular programme elements being evaluated and from the overall accumulated experience of the evaluator(s) in their fields of expertise.

4. Formulating Recommendations

Recommendations are actions for consideration by intended users of the evaluation results. Formulating recommendations is a distinct stage of programme evaluation that requires information beyond what is necessary to assess programme performance. For instance, knowing that a programme is able to reduce the risk of disease does not necessarily translate into a recommendation to continue the effort, particularly when competing priorities or other effective alternatives exist. Thus, recommendations for continuing, expanding, redesigning, or terminating a programme or project are separate from assessments regarding a programme's effectiveness. Making recommendations requires information on the context, particularly the organizational context in which programmatic decisions will be made, of which the evaluator(s) may not always be sufficiently informed.

The various elements of evaluation explained above are illustrated with information from a UNFPA project evaluation in **Box 1** below.

Box 1. An illustration of Evaluation Purpose, Objectives, Findings, Conclusions, Lessons and Recommendations: the Jenin Community-based RH Project, Palestine.

Evaluation purpose: to enable decision on continued funding of the project and provide information on changes required in the project approach.

Evaluation objectives: to assess the strengths and limitations of the project in terms of: design; management; quality of care; and cost effectiveness.

Data Collection Methods

1. Review of available documentation to obtain a general overview of objectives, design and logistics.
2. Interview with key contacts from organizations associated with the project.
3. Focus group discussions with Family Planning Educators and clients at service delivery points (SDPs)
4. Client satisfaction survey.
5. Visit to selected SDPs and application of checklists to assess the physical conditions of the facilities, inventory of contraceptive supplies, and observe the quality of service delivery.
6. Performance of patient record audit in order to determine if they were adequately completed and if essential indicators for diagnosis and treatment were provided.

Sample finding, conclusions, lesson, and recommendations regarding one aspect of quality of care (i.e. presence of female service providers).

Finding: interviews with CARE and PFS staff as well as focus group discussions with clients revealed that female physicians are acceptable to the women whereas male physicians are not.

The point is illustrated by the fact that utilization of clinic services increased only after the employment of a female physician. At the same time, the scarcity of female physicians in general, and in the Jenin District in particular, makes it hard to lure female physicians from the lucrative private practice. Investment in staff, both financial and in terms of training is vital for ensuring that experienced and skilled female physicians are involved in the project.

Detailed conclusion: increasing the availability of experienced female health professionals in the clinics based on the client load would improve the quality of care and eventually increase the number of clients.

Overall conclusions on quality of care: project activities have improved access to a range of culturally and technically appropriate and reasonably priced RH/FP services and information. However, there are gaps in quality of care including lack of basic sanitary facilities; poor clinic and staff management practices such as lack of supervision and support systems for the physicians [poor working environment].

.../

Box 1. (cont'd)

Lesson learned: in Palestine, as in many other Muslim Arab countries, female physicians, as opposed to male physicians, contribute to both the acceptability of services and continuing use of the clinics by female clients.

Overall recommendation: continue to support family planning services that are integrated with RH services using existing PHC services as a foundation. Within this context, continue support for the Jenin Community-based RH project.

Detailed recommendation: invest in high quality, culturally adequate service providers and clinic facilities: employ female physicians; ensure close and ongoing supervision of providers' performance by a project manager, preferably a female with a health background; upgrade the physical infrastructure of the clinics.

NOTE: UNFPA support to the project was not continued in the new CP cycle because USAID support was going to include the project area.

Source: Halabi 2000.

Sources

Centres for Disease Control and Prevention (CDC). "Steps in Program Evaluation", CDC Evaluation Working Group, September 1999. Available in English on the web at: <http://www.cdc.gov/eval/steps.htm>

Halabi, Hanan; Salem, Ruwaida; Wick, Laura. "Jenin Community Based RH Education Project", Project-end Evaluation for UNFPA. Birzeit University, Institute of Community and Public Health, January 2000.

Ministry of Foreign Affairs, Danida. "Evaluation Guidelines", February 1999.

UNICEF. "EVALUATION - A UNICEF Guide for Monitoring and Evaluation - Making a Difference?", Evaluation Office, 1991.

August 2004¹

PART II: DEFINING EVALUATION QUESTIONS AND MEASUREMENT STANDARDS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This part II of tool number 5 discusses the "what" of evaluation: steps to define evaluation questions and measurement standards. The content is based on a review of evaluation literature from bilateral and other development agencies such as such as Danida, ILO, Management Sciences for Health as well as documentation from UNFPA project evaluations.

II. Defining Evaluation Questions

Most evaluations are concerned with issues of programme design, delivery and performance. Design and delivery issues refer to factors affecting results. These factors appear during programme implementation.

Performance issues relate to the actual programme results (see Box 1). Each of these issues is explained in greater detail below.

Box 1. What do we mean by Result?

A result is a describable or measurable change in state that is derived from a cause and effect relationship. Results are the effects generated by a programme.

There are *three different types* of results:

- Outputs** Products and services that result from the completion of activities within a development intervention.
- Outcomes** The intended or achieved short and medium-term effects of an intervention's outputs, usually requiring the collective effort of partners. Outcomes represent changes in development conditions which occur between the completion of outputs and the achievement of impact.
- Impacts** Positive and negative long term effects on identifiable population groups produced by a development intervention, directly or indirectly, intended or unintended. These effects can be economic, socio-cultural, institutional, environmental, technological or of other types.

Source: Tool Number 1: Glossary of Planning, Monitoring and Evaluation Terms, March 2004.

¹ This tool was first published in December 2000.

Validity of design

A good programme design guides the implementation process, facilitates monitoring of implementation and provides a solid base for performance evaluation. In UNFPA, issues of programme design are assessed by using the programme logical framework.

Some key questions related to design include²:

- **Outputs, outcomes and impact (the results):** are they clearly stated, describing solutions to identified problems and needs?
- **Inputs and strategies:** are they identified and are they realistic, appropriate and adequate to achieve the results?
- **Indicators:** are they direct, objective, practical and adequate (DOPA)³? Is responsibility for tracking them clearly identified?
- **External factors and risks:** have factors external to the programme that could affect implementation been identified and have the assumptions about such risk factors been validated?
- **Execution, implementation, monitoring and evaluation responsibilities:** have they been clearly identified?
- **Gender sensitivity:** does the programme design address the prevailing gender situation? Are the expected gender related changes adequately described in the outputs? Are the identified gender indicators adequate?
- **Capacity building:** does the programme include strategies to promote national capacity building?
- **Programme approach:**
 1. In the case of a programme evaluation, does the design clearly establish linkages among programmes components?
 2. In the case of a programme component evaluation, are linkages among its interventions clearly established to ensure synergy in achievement of programme components results?

² These questions are illustrative and should not be used as a "blue print".

³ A **Direct** Indicator closely tracks the result it is intended to measure; an **Objective** Indicator is unambiguous about: 1) what is being measured and data being collected; 2) has a clear operational definition that is independent of the person measuring the indicator; a **Practical** Indicator can be gathered at reasonable cost and frequency, and can be available in time for use in decision-making; an **Adequate** indicator constitutes the minimum necessary to ensure that progress towards results is sufficiently well captured. Further details on indicators are provided in Tool Number 6 Part I: Identifying Output Indicators - The Basic Concepts available at www.unfpa.org

Delivery process

An assessment of the delivery process focuses on how the programme is being/was implemented to determine if the programme has remained on the right track towards the achievement of planned results and if not, what were the influencing factors.

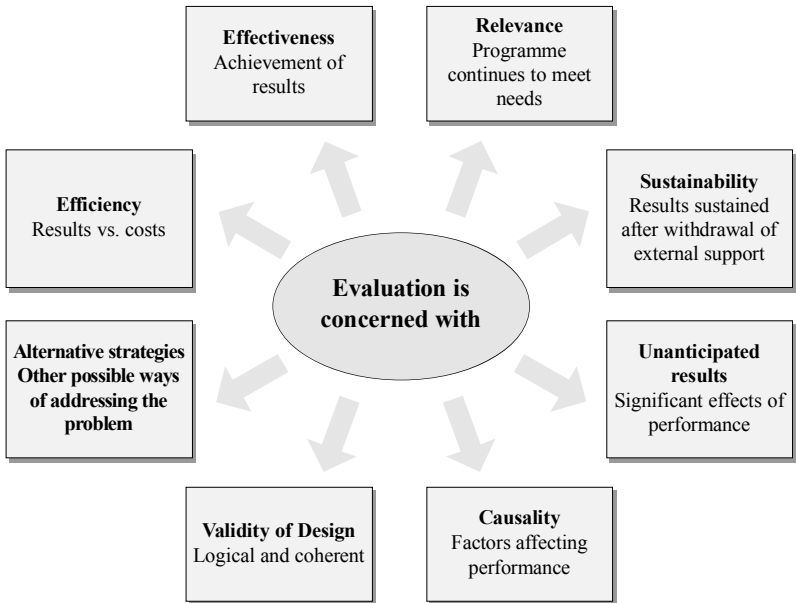
Some key questions related to the delivery process include:

- **Activities:** how were they implemented?
- **Outputs:** were the planned outputs achieved? Were they achieved within the planned time frame? Were they of adequate quality? If not, why?
- **Programme management:**
 1. Did the programme implementers discharge their roles in a cost-effective and cost-efficient manner? If not, why not?
 2. Were sound financial and equipment management procedures practised? Were the financial, human and material resources managed responsibly and efficiently?
 3. Was the technical assistance provided appropriate and of good quality?
 4. Did the monitoring and evaluation systems and processes allow for adequate assessment of changes in risks and opportunities in the internal and external environments? Did they contribute to effective decision-making in the course of programme implementation?

Performance

When assessing programme performance, evaluations look beyond the delivery process and focus on the results of inputs delivered and the work done. The outcome of this assessment determines whether or not the programme has achieved or is likely to achieve its outputs and contribute to achieving programme outcomes and impact.

The core evaluation concerns to assess programme performance are illustrated in **Figure 1** and described below.

Figure 1. Core Evaluation Objectives

Source: ILO, 1997.

Relevance

An assessment of programme relevance examines the appropriateness of results in relation to: the national needs, policies, and priorities; the needs and priorities of programme target groups (the local programme context); UNFPA's policies and priorities and its comparative advantage vis à vis other UN agencies and development partners. The analysis ascertains whether the programme continues to make sense and identifies any changes that may have occurred in its context during implementation. The initial problems and needs may no longer exist and policies and priorities may have changed as a result of political, economic, social and other factors, or even because of programme activities. Ultimately, the analysis determines whether the results are still valid or should be reformulated.

Some key questions related to relevance include:

- Needs, mandates, policies and priorities:** Do the programme planned results address the national needs? Are they in line with the government's priorities and policies? Are they in line with UNFPA's mandate? Does the target population consider them useful? Are they complementary to other donor interventions? Should results be adjusted, eliminated or new ones added in light of new needs, priorities and policies?

Effectiveness

An assessment of programme effectiveness focuses on the extent to which the outputs have been or will be achieved and whether the programme is likely to contribute to the stated outcomes and impact. If not, the evaluation will identify whether the results should be modified (in case of a mid-term evaluation) or the programme be extended (in case of a final evaluation) in order to enable achievement of stated results.

Some key questions related to effectiveness include:

- **Outputs:** to what extent have planned outputs been or will be achieved? What is the quality of the outputs?
- **Data on indicators:** have data been collected on the indicators of achievement? Do they provide adequate evidence regarding achievement of programme outputs and contribution to outcomes and impact? Is it necessary to collect additional data?
- **Gender:** what were the achievements in terms of promoting gender equity and equality (planned/unplanned)?
- **Capacity development:** what were the achievements in terms of capacity development (planned/unplanned)?

Efficiency

An assessment of programme efficiency measures the "productivity" of the programme interventions. It assesses the results obtained in relation to the expenditure incurred and resources used by the programme during a given period of time. The analysis focuses on the relationship between the quantity, quality, and timeliness of inputs, including personnel, consultants, travel, training, equipment and miscellaneous costs, and the quantity, quality, and timeliness of the outputs produced and delivered. It ascertains whether there was adequate justification for the expenditure incurred and examines whether the resources were spent as economically as possible.

Some key questions related to efficiency include:

- **Costs:** did the actual or expected outputs justify the costs incurred? Have the resources been spent as economically as possible?
- **Duplication:** did programme activities overlap and duplicate other similar interventions (funded nationally and/or by other donors)?
- **Alternative options:** are there more efficient ways and means of delivering more and better outputs with the available inputs?

Sustainability

An assessment of programme sustainability ascertains the extent to which the programme results have had or are likely to have lasting results after programme termination and the withdrawal of external resources. The factors affecting sustainability are examined on the basis of the priority assigned to the programme by stakeholders. Their readiness to continue supporting or carrying out specific activities, or even replicate the activities in other regions or sectors of the country, is particularly relevant. The analysis also assesses the availability of local management, financial and human resources that would be needed to maintain the programme results in the long run.

Some key questions related to sustainability include:

- **Likely sustainability:** is it likely that programme achievements will be sustained after the withdrawal of external support? Are involved counterparts willing and able to continue programme activities on their own? Have programme activities been integrated into current practices of counterpart institutions and/or the target population?
- **Resources:** have they been allocated by the counterparts to continue programme activities?

Causality

An assessment of causality examines the factors or events that have affected the programme results. If the inputs needed to carry out the planned activities and deliver the expected outputs were available on time, the implementation and performance would be successful. If, on the other hand, there were significant deviations from the planned schedules, the analysis would determine the reasons for such changes. The assessment should also analyse the effect of other factors such as technical, administrative or managerial constraints, inadequate inputs, failed commitment by programme counterparts, insufficient funds, faulty assumptions or the effect of unexpected external factors.

Some key questions related to causality include:

- **What factors:** what particular factors or events have affected the programme results?
- **Internal/external factors:** were these factors internal or external to the programme?

Unanticipated results

A programme evaluation may find significant unforeseen positive or negative results of programme activities. Once identified, appropriate action can be taken to enhance or mitigate them for a greater overall impact.

Some key questions related to unanticipated results include:

- **Were there any** unexpected positive and/or negative results of the programme?
- **How to address them:** can they be either enhanced or mitigated to achieve the desired impact?

Alternative strategies

Evaluations examine whether alternative approaches might have had greater impact or might have been more cost-effective, particularly if the original strategies turn out to be inappropriate. This analysis is especially valuable when follow-up programmes are planned.

Some key questions related to alternative strategies include:

- **More effective approaches:** is there, or would there have been, a more effective way of addressing the problem(s) and satisfying the needs in order to achieve the outputs and contribute to higher level aims?
- **Relevance:** are programme strategies still valid or should they be reformulated?

Not all of the above evaluation objectives have to be examined in every evaluation. The final choice will depend on the purpose of each evaluation⁴. For instance, a formative evaluation undertaken in the course of programme implementation with the aim of taking decisions to improve its design and/or implementation would typically emphasise concerns of design, delivery process, efficiency, causality, unanticipated results, and alternative strategies.

A summative evaluation, undertaken at the end of programme implementation to judge its effectiveness, would typically concentrate on concerns of relevance, effectiveness, efficiency, alternative strategies and sustainability.

An evaluation, which aims at extracting lessons learned and best practices or defining policy options would assess design, delivery processes, causality and efficiency in order to extract those characteristics which can effectively and efficiently deliver the desired results.

⁴ For a discussion on the issue of evaluation purpose, consult Tool Number 3: Purposes of Evaluation.

III. Methodological Challenges

Evaluator(s) face a number of methodological challenges with respect to the standards⁵ they use to measure relevance, effectiveness, efficiency and sustainability. These standards and methodological challenges are summarized in **Table 1** and further discussed below.

Table 1. Performance related Evaluation Objectives: measurement standards and methodological challenges.

Evaluation Objective	Measurement Standards	Methodological Challenge
Relevance	Needs, priorities and policies of programme target population, counterparts; UNFPA's policies, priorities, comparative advantage.	Lack of consensus on or incorrect assessment of needs and country priorities and lack of clear policies. Incorrect assessment of and/or lack of consensus on UNFPA's comparative advantage.
Effectiveness	Agreed outputs, outcomes and impact. Status of affected institutions, target population, and infrastructure prior to the programme interventions.	Unclear, multiple, confusing or changing results statements. Poorly defined results indicators. Lack of baseline information on the affected institutions, people, infrastructure. Poor knowledge of cause /effect linkages. Difficulty in attributing results to the particular programme due to intervening variables.
Efficiency	Similar interventions/best practices; criteria for what is considered reasonable.	What standards to use as a reference.
Sustainability	Sustainability factors (see Box 4)	Long term sustainability is a hypothetical, projected situation. Not all intervening factors which can compromise sustainability can be foreseen.

Source: Adapted from Danida, 1999.

⁵ A standard is a level of performance according to specified criteria or achievement dimensions.

With respect to the **relevance** of programme strategies, it requires in-depth field analysis by evaluator(s) to adequately determine the continued relevance of the planned programme results if the context (needs, priorities and policies) were not clearly defined at the time of programme design or in the course of implementation. **Box 2** provides an example of such a situation from a UNFPA commissioned evaluation.

Box 2. The importance of using contextual information for programme design and adjustment.

The report of an evaluation of the Jenin Community-Based RH project in Palestine concluded the following:

"The second objective stated in the project proposal was that "18,000 new users of family planning will have been recruited and continuation rates will be improved."

The problem with this objective is twofold. First of all, the target population was overestimated and the target recruitment numbers were too ambitious. The number of target population does not appear to be based on the available scientific evidence (estimates of the district population published by the Palestinian Central Bureau of Statistics). Secondly, the manner in which the numbers were determined is not clear from the proposal document. Furthermore, the project staff did not seem to be aware of this target and were unable to explain it. Still, the unrealistically high expectations seemed to put pressure on the staff to generate high numbers of new users and to reflect the numbers in their reported statistics, to the exclusion of other important activities that were being carried out.

As for the objective of improving continuation rates, it was not clear how that would be verified, since no baseline figure for continuation rates in those communities existed.

As a consequence of lack of proper initial situation analysis and adjustment in the course of project implementation, a large proportion of the budget (80,000 USD) was allotted to the purchase of contraceptives many of which eventually expired on the shelf due to lack of demand."

Source: Halabi, January 2000.

Standards to determine a programme's **effectiveness** often have to be reconstructed by evaluator(s) when planned programme results are too ambitious in relation to the resources and time frame available. Additionally, the programme indicators are frequently poorly or incorrectly defined thus hampering a sound assessment of achievement of programme outputs. The failure of programme implementers to gather baseline data at the beginning of the implementation process against which progress can be measured, constrains the evaluator(s) ability to assess results. To facilitate the objective evaluation of results achievement, the programme indicators, particularly those related to outputs, should be adjusted and refined in the early phase of programme implementation based on collection of baseline data and the accumulated knowledge of the programme context.

An additional difficulty is that long-term results can usually only be determined with certainty a significant period of time after programme completion. During that time, developments external to the programme such as economic and social development factors, (for instance increase in age at marriage) could have influenced the programme targets thereby making it difficult to ascribe improvements to the programme interventions.

Defining objective *efficiency* standards is a major challenge for evaluators of UNFPA's programmes and projects. In practice the evaluator(s) frequently rely on their expert judgment, which can be subjective. However, approaches are available to define standards, among others in the health field. The Continuous Quality Improvement tool to strengthen FP programmes is one such approach, which, if used in the course of programme implementation, greatly facilitates monitoring and evaluation of programme efficiency (see **Box 3**). Another good approach for identifying efficiency standards is "benchmarking", analysing the performance of organizations, which excel in areas of work relevant to UNFPA.

Box 3. Using the Continuous Quality Improvement (CQI) approach to define efficiency standards

CQI is performed by teams of staff at RH service delivery points. The CQI team implements the 7 step CQI cycle:

- Step 1: Identify an area** where opportunities for improvement exist
- Step 2: Define a problem** within that area, and **outline the sequence of activities (the process)** that occurs in that problem area
- Step 3: Establish the desired outcomes** of the process and the **requirements needed** to achieve them
- Step 4: Select specific steps** in the process to study and for each step, list the factors that prevent the achievement of the desired outcome
- Step 5: Collect and analyze data** about the factors that are preventing the achievement of the desired outcomes of the specific step being studied, and quantify the outcomes of that step
- Step 6: Take corrective action** to improve the process
- Step 7: Monitor the results** of the actions taken.

In **step 3** the CQI team **defines the standards** of efficiency against which services will be monitored and evaluated. The following is an example of such a standard:

"The client registration process is completed within 30 minutes of client's arrival at the clinic"

Source: The Family Planning Manager, Volume II, Number 1, January/February 1993.

Long-term programme **sustainability** is hard to foresee as many factors intervene over time. For instance, governments change and so may policies that are critical to support certain programmes originally funded by UNFPA. With a government change, key administrators also change and with them the institutional memory so necessary to keep particular approaches and programmes running. A severe economic crisis may appear, jeopardizing funding for the programme. However, programme designers must ensure that the sustainability factors listed in **Box 4** are fully considered at the time of situation analysis and programme design. Evaluator(s) assess the likelihood of sustaining programme activities using the same standards.

Box 4. Factors which influence sustainability of programme activities

1. **Policy Support Measures**, priorities and commitments of programme implementers and target groups.
2. **Choice of Technology** (for instance contraceptives) is appropriate to existing socio-cultural and economic conditions.
3. **Environmental aspects** such as management of population growth and distribution in relation to available land, water, fuel. Management of their living conditions such as housing, waste disposal, drinking water supply in order to avoid epidemics.
4. **Socio-cultural** integration. Acceptance of interventions because they are consistent with local traditions of groups (gender, ethnic, religious).
5. **Organizational** capacity to manage programme activities.
6. **Economic** viability and financial sustainability.

Source: Danida, 1999.

Sources

Halabi, Hanan; Salem, Ruwaida; Wick, Laura. "**Jenin Community Based RH Education Project**", Project-end Evaluation for UNFPA, Birzeit University, Institute of Community and Public Health, January 2000.

ILO. "**Guidelines for the Preparation of Independent Evaluations of ILO Programmes and Projects**", Evaluation Unit, Bureau of Programming and Management, November 1997. Available in English, French and Spanish on the web at:

<http://www.ilo.org/public/english/bureau/program/guides/evalmenu.htm>

Management Sciences for Health. "**The Family Planning Manager**", Volume II, Number 1, January/February 1993. The full Health Manager's Toolkit is available in English, French and Spanish on the web at <http://erc.msh.org/index.cfm>

Ministry of Foreign Affairs, Danida. "**Evaluation Guidelines**", February 1999.

August 2004¹

PART III: THE DATA COLLECTION PROCESS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

Part III of tool number 5 discusses the "how" of programme evaluation, namely the data collection process, including determination of data collection methods, data analysis and interpretation. The content is based on a review of evaluation and other literature from bilateral and other development agencies such as Danida, Management Sciences for Health, Save the Children, UNFPA, UNICEF, USAID, and the W.V. Kellogg Foundation.

II. Determining information needs

Once the evaluation purposes, objectives, questions and standards for assessing the programme have been selected², **the adequacy of existing information** to provide answers to the evaluation questions which meet the selected measurement standards **should be reviewed**. Up to date programme results statements (outputs, outcomes and impact) and corresponding indicators as stated in the programme results and resources frameworks (RRFs) are some of the readily available information on standards established for UNFPA's programmes and their components. It is important to ensure that indicators and means of verification are regularly reviewed and updated to enable effective assessment of programme progress and performance.

Additional information to be used by the evaluation, including analysis of implementation processes to achieve planned results, can be obtained from programme work plans; progress and evaluation reports; field monitoring visit reports; technical assessments and survey reports; clinic statistics; research reports; government policy documents and the like. Analysis of existing data can be helpful to refine evaluation questions, identify informants for subsequent interviewing, develop interview protocols, and determine what data important to the evaluation is missing and should be collected by the evaluator(s). **Box 1** highlights some useful criteria for determining the need for additional data.

¹ This tool was first published in May 2001.

² For a detailed discussion of these concepts, see Tool Number 3: Purposes of Evaluation and Tool Number 5, Part II: Defining Evaluation Questions and Measurement Standards.

Box 1. Useful questions to help determine the need for additional data.

- What level of detail is required? What difference would it make if additional information is or is not obtained?
- How will the additional information be used? It is important to collect only the information, which will be used and to use all the information collected.
- How credible are different types of data to the intended users of evaluation results? The level of credibility of data sources and data collection methods determines the acceptance and use of evaluation conclusions and recommendations by the intended users.
- When is the information needed? Time constraints may determine the length and nature of additional data collection exercises.
- What resources are available for the evaluation? The availability of expertise and financial resources determines the sophistication of additional data collection.

Source: Adapted from UNICEF, 1991.

III. Determining methods for collecting additional data

The next step is to identify how to collect the additional data required. **Quantitative and qualitative data collection methods** as well as **deductive and inductive analytical approaches** can be used for this purpose.

Quantitative and Qualitative data collection methods include:

- questioning people through individual and group interviews such as focus group discussions and community interviews;
- conducting surveys;
- observing people, processes, objects, conditions, and situations.

Annex 1 further describes data collection methods.

Quantitative and qualitative data collection methods each have their strengths and weaknesses and are suited to answer different types of questions as highlighted in **Table 1**.

Table 1. Characteristics of Quantitative and Qualitative Data Collection

	Quantitative Methods	Qualitative Methods
Use	To numerically measure "who, what, when, where, how much, how many, how often"	To qualitatively analyse "how and why"
Examples	Standardized interviews; surveys using closed-ended questions; observation.	Free and guided interviews (including focus group); surveys using open-ended questions; observation; interpretation of documents.
Strengths	<ul style="list-style-type: none"> ■ Provide quantitative, accurate and precise "hard data" to prove that certain problems exist ■ Can test statistical relationships between a problem and apparent causes ■ Can provide a broad view of a whole population ■ Enable comparisons ■ Establish baseline information which can be used for evaluating impact. 	<ul style="list-style-type: none"> ■ Useful when planning a programme concerned with social change ■ Provide a thorough understanding of programme/project context in order to interpret quantitative data ■ Provide insights into attitudes, beliefs, motives and behaviours of a small sample population (families, communities) ■ Establish baseline information which can be used for evaluating qualitative outcomes (changes in knowledge, attitudes, behaviours, institutional processes etc.) ■ Useful in case of money and time constraints ■ Useful for getting feed-back from stakeholders.
Weaknesses	<ul style="list-style-type: none"> ■ May be precise but not measure what is intended ■ Cannot explain the underlying causes of situations. 	<ul style="list-style-type: none"> ■ Are generally not representative; do not allow generalizations ■ Susceptible to biases of interviewers, observers and informants.

Sources: UNICEF, 1991; Gosling, 1995. USAID TIPS Number 2, 1996.

In answering evaluation questions, to avoid dependence on the validity of any one source, it is highly recommended to use a combination of different types of qualitative and quantitative data collection methods such as review of statistics, small-scale surveys, interviews and observation (also called **data triangulation**). Observation is an invaluable methodology to collect data that surveys and interviews cannot accurately capture. For instance, observation is necessary to assess client/provider or student/teacher interaction on sensitive subjects such as sexual and reproductive practices. Relying only on surveys and interviews in this situation may not yield accurate information as respondents tend to report ideal not actual behaviours.

Similarly, carefully study of various materials produced by the programme such as IEC materials (on sexuality, HIV/AIDS prevention etc.), training modules, policies, and guidelines, can provide valuable information and insights on how the issues are tackled. For example, by reviewing IEC materials, an evaluation of a UNFPA funded HIV/AIDS prevention project found that brochures designed to increase awareness on ways to avoid becoming HIV infected did not mention condoms but recommended that "people not go dancing in places where one can catch HIV/AIDS!"

Finally, quantitative surveys do not enable exploration of underlying causes. Thus, a combination of methods provides a more complete analysis of the subject matter being evaluated thereby enhancing the credibility of the evaluation conclusions and recommendations. **Box 2** summarizes a few criteria to guide selection of methods to collect additional evaluation data.

Box 2. Criteria for selecting data collection methods

- Determine which data-collection methods best answer key evaluation questions.
- Tie method selection to available resources. This may mean revising the evaluation design and methods, or determining other options to stay within budget. It may also mean finding additional resources to fund the most effective and useful evaluation design.
- Choose methods, which will facilitate the participation of key programme stakeholders in the evaluation.
- Strengthen the credibility and usefulness of evaluation results by mixing evaluation methods where appropriate.

Source: W.K. Kellogg Foundation, 1998.

Evaluator(s) are not only concerned with what data collection methods to use in order to adequately address evaluation objectives and question. They also need to select a certain analytical approach to gathering information. When using a **deductive** approach, evaluator(s) formulate hypothetical answers to the evaluation questions at an early stage of the evaluation process based on available information and the evaluator(s) accumulated knowledge of the subject matter being evaluated. Data is then collected to confirm or refute these hypotheses. When using an **inductive** approach, the evaluator(s) start with an open, questioning mind. They gradually uncover issues and themes through iterative field observation, interviews and data analysis thus leading to a deeper understanding of the subject matter.

While most evaluations rely on a combination of the two approaches, a deductive approach would be suitable for addressing evaluation objectives of efficiency and effectiveness. A deductive approach would, for instance, be used to examine whether the best results were achieved with the inputs provided and activities implemented and whether the planned results were achieved. An inductive approach would be very useful for addressing evaluation objectives of relevance, impact and sustainability. It is particularly useful for evaluating socio-cultural aspects of a programme because there is limited knowledge about the cause-effect relationships among programme inputs, outputs and outcomes.

IV. Analyzing and Interpreting Data

The evaluation information collected must be described, analyzed, interpreted, and a judgment made about the meaning of the findings in the programme context. Interpretation involves looking beyond the raw data to ask questions about what they mean, what the most significant findings are, and what conclusions and recommendations should be drawn from these findings. A few basic techniques for organizing and analyzing data are described below.

Quantitative Analysis

Quantitative data analysis interprets the numerical findings considering the programme context. As implementers of programme activities are most knowledgeable about the context, they should work together with the evaluator(s) to assess whether the figures make sense; whether they adequately reflect programme results; what possible explanations are for unexpected figures; what conclusions and recommendations can be drawn from the figures.

Qualitative Data Analysis

While some accounts resulting from in-depth interviews and focus group discussions are stand-alone illustrations of important themes of the evaluation, it is, in most cases, valuable to analyze qualitative data more systematically.

Analysis of qualitative data from interview transcripts, observation field notes or open-ended surveys can identify similarities across several accounts, as well as directions, trends and tendencies. Data can be categorized into recurrent themes and topics that seem relevant to answer the evaluation questions and to develop new or test already selected hypotheses.

However, evaluators run the risk of drawing hasty conclusions and making generalizations when breaking transcripts and field notes up into thematic categories. They can avoid this problem by writing case studies and narrative summaries, which highlight the context and particular characteristics of key pieces of the programme being evaluated.

Another problem frequently encountered when analyzing qualitative data is the excessive focus on "quantifying" qualitative data and interpreting qualitative data as if it were quantitative data. For example, when analyzing and interpreting focus group discussion data, some evaluators tend to tabulate the responses and report on them in terms of ratios and percentages rather than exploring further the information, ideas, opinions and attitudes which can help answer the evaluation questions "why?" and "how?"

Values and Biases

Biases and values inevitably influence both quantitative and qualitative data analysis. Evaluator(s) control for biases and values by triangulating multiple types of data (quantitative and qualitative), data collection methods, data sources, and perspectives or theories to interpret the data for instance

by involving stakeholders in analyzing and interpreting the data. Evaluator(s) can also ensure that they pay attention to the influences of biases and values through an ongoing process of writing descriptive memos about the evaluation process, their data, and their interpretations.

Annex 1. Data Collection Methods

The following list and description of data collection methods is not intended to be exhaustive. It is rather an overview of the key characteristics of the most commonly used data collection methods. These may be applied not only for evaluations but also at other stages of the programme cycle such as situation analysis, programme design, monitoring and reviews. Each method may be explored further in the referred sources.

A. Review of existing programme and other documents.

1. Programme specific information such as reports of programme progress, field monitoring visits, programme reviews, surveys, research and evaluations.

Such documents enable the evaluator to learn about the history, context, results, and outcomes of a particular programme. They also provide clues about important shifts in programme development and implementation. A document review may also be a good way to formulate questions for use in a survey or interview.

2. Other information not directly related to the programme such as research studies; government data such as clinic based statistics; and evaluations of similar programmes and projects. Evaluation databases such as the UNDP CEDAB and IFAD EKSYSY³ are good sources for increasing knowledge of lessons learned on issues which are present in all development programmes such as gender, capacity-building, and collaboration with NGOs.

It should be noted that written documents do not necessarily provide comprehensive or correct answers to specific problems, as they may contain errors, omissions, or exaggerations. They are simply one form of evidence, and should be used carefully and together with other types of data.

B. Questioning People.

1. **Interviews such as Key Informant, Focus Group Discussion and Community Interviews, and Nominal Group Technique.**

General Characteristics

Interviews, together with document reviews, are the most frequently used data collection method in UNFPA evaluations. Unstructured and guided interviews yield qualitative data. In unstructured

³ Further information on these databases can be found on the IFAD and UNDP evaluation web sites at http://www.ifad.org/list_eval.asp and <http://www.undp.org/oo/index.htm> respectively.

interviews, the interviewer's only guide are the evaluation objectives. Unstructured interviews are a good tool for exploring the opinions of respondents and uncovering unexpected factors. In a guided interview, the respondent is asked to provide information about items on a prepared checklist.

Standardized interviews yield quantitative data. They use a questionnaire with a fixed number of questions and sometimes a pre-selected range of possible answers.

In general, the more open-ended the interview the more deeply the respondents' feelings and perspectives can be understood; the more structured the interview, the more comparable the data. Many reports based on questionnaires provide an array of facts (percentages, breakdowns) but shed little light on people's motivations that could be built on to improve practices.

One of the first steps in interviewing is to identify knowledgeable informants, people who can provide pertinent and reliable information. Informants can be clients at service delivery points, programme implementing partners, community members, local leaders, politicians, or health professionals. Depending on the type of information needed, informants can be interviewed individually or in groups.

In-depth Interview

If the evaluator(s) are concerned about maintaining the informants' anonymity or simply want to make sure that they feel free to express controversial ideas, it is best to interview informants individually. This also allows the evaluator(s) to compare various perspectives of an event, which is particularly useful when exploring sensitive topics.

A key informant interview is a form of in-depth interview often used. Key informants are selected for their first-hand knowledge about the topic of interest. For example, the head of an HIV epidemiology unit may act as a key informant on information relating to the incidence of HIV/AIDS. Traditional birth attendants would be key informants for information on traditional management of pregnancy and delivery³.

Group Discussion

When confidentiality is not a concern, and the evaluator(s) are interested in quickly sampling a range of opinions on a topic, a group discussion is preferable. There are several types of group discussions. Focus group discussions, community and other types of group interviews are among those frequently used.

A Focus group discussion is an inexpensive, rapid appraisal technique through which a facilitator guides 7-10 people in a discussion of their experiences, feelings and preferences about a topic. The facilitator raises issues identified in a discussion guide and uses probing techniques to animate the discussion and promote in-depth reflection among focus group participants. Sessions typically last one to two hours. The facilitator's discussion guide should contain few items thereby allowing

³ For a stepwise explanation on how to conduct key informant interviews, consult TIPS Number 2, 1996 at http://www.dec.org/usaid_eval/#004

some time and flexibility to pursue unanticipated but relevant issues. In order to maximize exchanges among focus group participants they should share certain common characteristics, i.e. be of same sex, age group, and social background and have similar concerns. Many participants in focus group discussions find the interaction stimulating and mention things they would not have thought of individually⁵.

In community interviews, which usually take the form of community meetings open to all, interaction is between the interviewer and the participants. Such meetings are susceptible to manipulation from the more powerful members of the community and are less suitable to discuss sensitive issues.

Nominal Group Technique

In this technique, from five to seven people are asked by a leader to generate ideas on a single topic. Through discussion, a consensus is reached on a list of most important ideas. A single session, which deals with a single question, usually takes about 60-90 minutes. The nominal group technique was developed to facilitate efficient group decision-making by busy private sector executives. It may also be useful in evaluation, particularly when groups composed of experts, community members, or programme staff are making recommendations for ongoing programmes.

Box 3. To ensure reliability, validity and avoid bias when questioning people:

- Pre-test interview guides and questionnaires;
- Ensure that the group of key informants selected include all the groups which can provide information of significance for the evaluation;
- Assess the reliability of informants. Their knowledge, credibility, impartiality, willingness to respond, and the presence of outsiders who may inhibit their responses are important factors;
- Check investigator bias, including tendencies to concentrate on information that confirms preconceived notions and hypotheses;
- Be systematic in note taking by recording the exact words, facial and body expressions descriptively rather than analytically, and trying not to let own perceptions of what is being said and expressed interfere while recording;
- Check for evidence that calls into question preliminary findings and thus bring out issues which may have been overlooked;
- Get feed-back from informants on major findings.

Source: W. K Kellogg Foundation, 1998.

⁵ For step-wise guidance on how to conduct focus group discussions consult TIPS Number 10 at http://www.dec.org/usaid_eval/#004

2. Surveys

There are several types of surveys:

Censuses: a complete enumeration of all units in a population.

Formal large-scale sample surveys (for instance DHS surveys): a randomly drawn representative sub-group from which researchers generalize about the whole population.

Informal small-scale sample surveys (for instance KAP surveys): a small non-random (such as purposeful selection of people in different categories on the basis of easy accessibility) sample of 30-50 individuals who are asked a few questions (10-20).

Large-scale technically complex surveys should be avoided in programme evaluation as they are expensive and time-consuming. Informal, small-scale sample surveys can, however, provide useful quantitative data, for instance, on use of and access to RH services to complement other qualitative evaluation data.

Survey questions can be open-ended or closed-ended. Open-ended questions might ask: How do you feel about the program? What do you want to see happen in your community? Open-ended questions provide relatively rich information about a topic and allow participants to report thoughts, opinions and feelings. However, there are disadvantages. Sometimes people are reluctant to express opinions, or the survey may be time-consuming to complete and analyze.

Unlike open-ended questions, closed-ended questions provide discrete, multiple-choice responses from which the respondent selects the most appropriate answer. For example:

How often do you use our center?

- a. never
- b. a few times a year
- c. once a month
- d. a few times a month
- e. once a week
- f. more than once a week

Closed-ended questions have the advantage of uniformity and easy translation for statistical analyses. Surveys with closed-ended questions can easily be administered to large groups of people and are usually easy to complete. However, they tend to impose a set of fixed ideas or values on the respondent by forcing choices from a limited array of options. As a result, they are less likely to uncover new and unexpected information, and they limit the emergence of in-depth understandings and nuances of meanings. In general, written survey questions are inappropriate if the respondents have low literacy or are unfamiliar with the conventions of survey completion. A survey administered in person might be more appropriate for this population.

A survey is only as good as the people administering it, so care should be given to selecting, training and supervising surveyors.

C. Observation

Evaluator(s) record what they see and hear at the project site using an observation checklist. Observation may be of physical surroundings, ongoing activities, processes and discussions.

Observation may be useful:

- When performance monitoring data indicate that results are not being accomplished as planned, and when implementation problems are suspected, but not understood. Direct observation can help identify whether the process is poorly implemented or required inputs are absent;
- When details of an activity's process need to be assessed, such as whether tasks are being implemented according to standards required;
- When an inventory of physical facilities and inputs is needed and not available from existing sources;
- When interview methods are unlikely to elicit needed information accurately or reliably, either because the respondents don't know or may be reluctant to say;
- To formulate questions which can be asked in subsequent interviews.

It is important to distinguish between observation and interpretation of what is seen. An evaluator should also recognize that even the most passive, unobtrusive observer is likely to affect the events under observation. Just because you observe it, do not assume that you are witnessing an event in its "natural" state⁶.

D. Rapid Appraisal

Rapid appraisal is essentially the use of a mix of the above-described methods in order for decision-makers to obtain timely, relevant, accurate and usable information on development programmes and projects. Key informant, focus group, and community interviews, observation and informal surveys are the methods most commonly used by rapid appraisal⁷.

⁶ For useful guidance on how to improve the quality of direct observation, consult TIPS Number 4, 1996 at http://www.dec.org/usaaid_eval/#004

⁷ An example of a rapid appraisal methodology used by UNFPA to assess national execution capacity is described in Evaluation Findings Issue 29, March 2000 available in English, French and Spanish at <http://www.unfpa.org/monitoring/reports.htm>. UNFPA Country Offices have also developed rapid appraisal methodologies for assessing the quality of RH service delivery.

Sources

Gosling, Luisa; Edwards, Mike. **"Toolkits - A Practical Guide to Assessment, Monitoring, Review and Evaluation"**, Development Manual 5, Save the Children, 1995.

Ministry of Foreign Affairs, Danida. **"Evaluation Guidelines"**, February 1999.

Patton, Michael Quinn. **"Qualitative Evaluation Checklist"**, September 2003. Available online in English at <http://www.wmich.edu/evalctr/checklists/qec/index.htm>

UNFPA. **"Population and Reproductive Health Programmes: Applying Rapid Anthropological Assessment Procedures"**, Technical Report Number 39, 1997.

UNFPA. **"Readings in Population Research Methodology - Volume 6, Advanced Basic Tools"**, Social Development Center, Chicago, Illinois. 1993.

UNICEF. **"EVALUATION - A UNICEF Guide for Monitoring and Evaluation - Making a Difference?"**, Evaluation Office, 1991.

USAID. Center for Development Information and Evaluation. Performance Monitoring and Evaluation TIPS. Available online in English at http://www.dec.org/usaaid_eval/#004

"Conducting Key informant Interview", 1996, Number 2.

"Using Direct Observation Techniques", 1996, Number 4.

"Using Rapid Appraisal Methods", 1996, Number 5.

"Conducting Focus Group Interviews", 1996, Number 10.

W. K. Kellogg Foundation. **"Evaluation Handbook"**, 1998.

PART IV: MANAGING THE EVALUATION PROCESS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This part IV of Tool number 5 discusses various aspects of managing the evaluation process: "who" does "what"; steps in the development of a terms of reference and in the selection of an evaluator/evaluation team; and pointers on managing and supervising the conduct of an evaluation. The tool provides overall guidance for a traditional approach to evaluation with limited stakeholder participation. However, the principles and management responsibilities mentioned in the tool should, with some adaptation, be applied to all types of evaluations. For further details on participatory monitoring and evaluation approaches, consult Tool Number 4: Stakeholder Participation in Monitoring and Evaluation.

The content is based on a review of the literature both from academia and international development agencies such as UNDP, UNICEF, WFP, Save the Children and bilateral donor agencies such as DANIDA, OECD, USAID and the U.S. Department of health and Human Services.

II. Defining the evaluation questions

As discussed in Tool number 3, there are different evaluation purposes and types of questions they can address. Part I and II of Tool number 5 discuss the steps involved in defining the evaluation objectives and questions to be answered by the evaluation. Once these have been established, the evaluation manager needs to ensure that the evaluation is carried out in a systematic and structured manner by following a few basic steps as outlined below.

¹ This tool was first published in May 2001.

III. Who does what: delineation of roles and responsibilities

People can participate in an evaluation in various capacities, as managers, as evaluators, or providers of information. It is essential, however, to have a clear delineation of roles and responsibilities among the various interested parties. **Boxes 1 and 2** highlight the potential responsibilities of the evaluation manager² and the evaluator(s).

Stakeholders are important partners in any evaluation and should be involved to varying degrees in the evaluation process. It is important to identify what roles the stakeholders will be expected to play in the evaluation and indicate the extent of their participation at the different stages of the evaluation process - which will vary with the type of evaluation carried out. When stakeholders are to be involved in all stages of the evaluation process (i.e., design and planning, information collection, the development of findings, evaluation reporting and results dissemination) then the evaluator's function might range from team leader to that of facilitator/resource person to be called on as needed. Stakeholder participation can facilitate evaluation capacity development and increase the likelihood of their acting on evaluation recommendations.

² The evaluation manager is usually a UNFPA staff member. Assigning a team composed of UNFPA staff to manage the evaluation can be useful to see the evaluation process through and ensure a higher quality product.

Box 1. Potential responsibilities of the evaluation manager**Preparation:**

- Determine the purpose and users of evaluation result
- Determine who needs to be involved in the evaluation process
- Define evaluation objectives and questions together with key stakeholders
- Draft the Terms of Reference (ToR) for the evaluation; indicate a reasonable time-frame for the evaluation
- Identify the mix of skills and experiences required in the evaluation team
- Oversee the collection of existing information/data; be selective and ensure that existing sources of information/data are reliable and of sufficiently high quality to yield meaningful evaluation results; information gathered should be manageable
- Commission/supervise the preparation of background document(s) as necessary
- Select, recruit and brief the evaluator(s)
- Ensure that background documentation/materials compiled are submitted to the evaluator(s) well in advance of the evaluation exercise so that the evaluator(s) have time to digest the materials
- Decide whose views should be sought (e.g., service providers, service users, central and/or local government counterparts, etc.)
- Develop additional information collection procedures and instruments (unless the evaluator is contracted to design information collection methods); ensure the use of a variety of data gathering methods to enhance the validity and credibility of evaluation results
- Propose an evaluation field visit plan
- Ensure availability of funds to carry out the evaluation
- Brief the evaluator(s) on the purpose of the evaluation; use this opportunity to go over documentation and review the evaluation work plan.

Implementation:

- Ensure that the evaluator(s) have full access to files, reports, publications and any other relevant information
- Ensure adequate administrative and logistical support during the evaluation
- Follow the progress of the evaluation; provide feedback and guidance to the evaluator(s) throughout all phases of implementation
- Assess the quality of the evaluation report(s) and discuss strengths and limitations with the evaluator(s) to ensure that the draft report satisfies the ToR, and that evaluation findings are defensible and recommendations are realistic
- Arrange for a meeting with the evaluator(s) and key stakeholders to discuss and comment on the draft report
- Approve the end product; ensure presentation of evaluation results to stakeholders; include country office programme staff in debriefing to promote information sharing and use of evaluation results.

Follow-up:

- Evaluate the performance of evaluator(s) and place it on record
- Disseminate evaluation results to the key stakeholders and other audiences (see Tool 5 part V)
- Promote the implementation of recommendations and use of evaluation results in present and future programming; monitor regularly to ensure that recommendations are acted upon.

Box 2. Potential responsibilities of the evaluator(s)

- Provide inputs regarding evaluation design; bring refinements and specificity to the evaluation objectives and questions
- Conduct the evaluation; as team leader supervise team members and manage the day-to-day process of carrying out the evaluation; make sure all aspects of the evaluation are covered
- Review information/documentation made available by the country office
- Design/refine instruments to collect additional information as needed; conduct or coordinate additional information gathering
- Undertake site visits; conduct interviews
- In the case of a participatory evaluation, facilitate stakeholder participation
- Provide regular progress reporting/briefing to the UNFPA evaluation manager
- As team leader act as mediator if there are dissenting views within the evaluation team
- Analyze and synthesize information; interpret findings, develop and discuss conclusions and recommendations; draw lessons learned
- Participate in discussions of the draft evaluation report; correct or rectify any factual errors or misinterpretations
- Guide reflection/discussions if expected to facilitate a presentation of evaluation findings in a seminar/workshop setting
- Finalize the evaluation report and prepare a presentation of evaluation results.

IV. Developing Terms of Reference for the evaluation

After the roles and responsibilities for implementing the evaluation have been delineated, the Terms of Reference (ToR) that lists the evaluation background as well as services and products the evaluator(s) is/are expected to deliver should be drafted. Discuss the ToR with the evaluator(s) and make any adjustments before initiating the evaluation. The ToR should:

- Provide background information on the history and current status of the programme/project being evaluated, including how it works (its objectives, strategies and management process), duration, budget and important stakeholders such as donors, partners, implementing organisations;
- Describe the purpose of the evaluation and who are its stakeholders; specify why the evaluation is being requested and what the information will be used for;
- Define the evaluation scope and focus. In consultation with stakeholders, identify the major evaluation objectives and questions in accordance with evaluation criteria such as: relevance, validity of design, effectiveness, efficiency, sustainability, impact, factors affecting performance, alternative strategies and unanticipated results (see Tool number 5, part II: Defining Evaluation Questions and Measurement Standards);

- Specify the evaluation methodology. Describe the data gathering instruments and methods of analysis. The methodology may be developed with the assistance of the evaluator(s);
- List the relevant information sources to be used by the evaluation such as monitoring, review, evaluation and other reports;
- Specify the composition of the evaluation team (e.g., number of team members, specify individual members' profile). The evaluation focus and methods as well as the availability of funds will determine the evaluation team composition. Multi-disciplinary teams, including specialists in UNFPA mandated substantive areas and at least one evaluation specialist, are often used to conduct evaluations of large programmes. Indicate who on the evaluation team will be the team leader;
- Specify the involvement of key stakeholders such as internal staff, programme partners, donors, and other stakeholders who will use evaluation results for decision-making. Detail the roles that each of these will play (see also Tool number 4: Stakeholder Participation in Monitoring and Evaluation);
- Describe the evaluation work plan. Specify the roles and responsibilities of the UNFPA evaluation manager, the evaluator(s) and the team leader; detail specific tasks to be undertaken as well as the time lines involved. Indicate which audiences are to receive which information at what times, what the nature and schedule of written reports and oral briefings will be, and how the findings will be disseminated and to whom;
- Specify logistics support required such as transportation, administrative support, translations, data processing, office and other equipment etc.;
- Specify the detailed evaluation budget including cost of consultants, travel, logistics, and support staff.

Table 1 is a sample evaluation plan format that provides an overview of the evaluation process. **Table 2** is a sample evaluation work plan. These are useful tools to assist the evaluation manager and team in managing the different levels of the evaluation process.

Table 1. The Evaluation Plan

Evaluation Objectives and Questions	Sources of Information	Location	Data Collection Methods	Responsible Party

Table 2. The Evaluation Work Plan

Tasks	Responsible Party	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct

V. Selecting appropriate evaluator(s)

The choice of evaluator(s) is an important factor in the effectiveness of evaluations. Evaluations can be conducted by **internal** or **external evaluators** or, as is often the case, by a combination. Careful consideration of the purpose of the evaluation will help to determine the best approach. Internal evaluations are conducted by evaluator(s) who is/are associated with the programme to be evaluated; external evaluations are conducted by evaluator(s) who is/are not associated with the execution, implementation and funding of the object of the study. For instance, if the purpose of the evaluation is to judge the overall effectiveness or impact of a programme then external evaluator(s) may be the better option given that they would not have a stake in the evaluation's findings, and that the results may be perceived as more objective. **Table 3** summarizes the possible advantages and disadvantages of using internal and external evaluators. A well-balanced combination of internal and external evaluators may be preferable for many purposes.

The inclusion of **national consultants** is useful and can enrich the evaluation exercise. They understand the evaluation context, and may provide the evaluation team with access to officials and sources of information that otherwise may not be available. Moreover, the inclusion of national consultants on a team can act as a catalyst for greater local "buy-in" into the evaluation results.

Table 3. Trade-Offs between Internal and External Evaluators

<i>Someone associated with the programme</i>	
<i>Advantages</i>	<i>Disadvantages</i>
<ul style="list-style-type: none"> ■ Knows the implementing organization, its programme and operations. ■ Understands and can interpret behavior and attitudes of members of the organization. ■ May possess important informal information. ■ Is known to staff, so may pose less threat of <ul style="list-style-type: none"> ■ anxiety or disruption. ■ Can more easily accept and promote use of evaluation results. ■ Is often less costly. ■ Doesn't require time-consuming recruitment negotiations. ■ Contributes to strengthening national evaluation capability. 	<ul style="list-style-type: none"> ■ May lack objectivity and thus reduce credibility of findings. ■ Tends to accept the position of the organization. ■ Is usually too busy to participate fully. ■ Is part of the authority structure and may be constrained by organizational role conflict. ■ May not be sufficiently knowledgeable or experienced to design and implement an evaluation. ■ May not have special subject matter expertise.
<i>Someone associated with the programme</i>	
<i>Advantages</i>	<i>Disadvantages</i>
<ul style="list-style-type: none"> ■ May be more objective and find it easier to formulate recommendations. ■ May be free from organizational bias. ■ May offer new perspective and additional insights. ■ May have greater evaluation skills and expertise in conducting an evaluation. ■ May provide greater technical expertise. ■ Able to dedicate him/herself full time to the evaluation. ■ Can serve as an arbitrator or facilitator between parties. ■ Can bring the organization into contact with <ul style="list-style-type: none"> ■ Additional technical resources. 	<ul style="list-style-type: none"> ■ May not know the organization, its policies, procedures, and personalities. ■ May be ignorant of constraints affecting feasibility of recommendations. ■ May be unfamiliar with the local political, cultural and economic environment. ■ May tend to produce overly theoretical evaluation results (if an academic institution is contracted). ■ May be perceived as an adversary arousing unnecessary anxiety. ■ May be costly. ■ Requires more time for contract, negotiations, orientation, and monitoring.

Source: Adapted from UNICEF Guide for Monitoring and Evaluation, 1991.

The evaluation purpose, methods and resources available will not only determine whom to select but also how many evaluators to recruit. In selecting candidates of **an evaluation team**, consider what each member will contribute to the evaluation not only in terms of his/her individual expertise and experience, but also in terms of his/her ability to function as member of a team. **Box 3** suggests various requirements to bear in mind when assembling an evaluation team.

Box 3. What should UNFPA look for in assembling a team?

- ✓ Evaluation skills (e.g. knowledge and practical application of evaluation methodologies)
- ✓ Knowledge of the subject region or country
- ✓ Subject matter expertise in the relevant area of UNFPA's mandate
- ✓ Analytical skills
- ✓ Facilitation skills in the event that participatory evaluation is undertaken
- ✓ Familiarity with UNFPA
- ✓ Team leadership skills
- ✓ Language proficiency
- ✓ Good drafting skills
- ✓ Demonstrated performance levels (check references)

VI. Managing and supervising the evaluation

Briefing evaluator(s)

At the beginning of the evaluation, the evaluation manager should meet with the evaluator(s) to ensure an understanding of the programme context, the evaluation purpose and approach, review the data collection instruments and the schedule of evaluation activities, and answer questions. Arrangements for administrative and logistical support should also be reviewed and any adjustments made. It is best at this juncture to discuss the format and content of the final evaluation report and the interim reporting arrangements.

Backstopping and supervising the evaluator(s)

The evaluation manager should consult with the evaluator(s) or the team leader at various times during the evaluation and request a debriefing at its conclusion; obtain feedback from the evaluator(s) on whether the evaluation instruments are appropriate for the purpose of the evaluation; comment on and assess the quality of draft reports; respect the independence of the evaluation and be prepared to accept findings and conclusions that may not support preconceived notions about the programme.

Finalizing the evaluation report

A report is needed to communicate evaluation findings, conclusions and recommendations (see Part I of Tool number 5 for further details and definitions of these evaluation elements). **Annex 1**

provides a suggested outline for the evaluation report. The report should be relatively brief and concise. It should either be written in or translated into the official language of the country. During the drafting process the evaluation manager should provide feedback and review the quality of the evaluation results. A draft report should be given for review to Government counterparts, programme managers and other users as appropriate. The evaluation manager and users of the report should discuss the findings, conclusions and recommendations, and provide comments on the draft report to the evaluator(s) before it is finalized. Such discussions can bring out new perspectives on the meaning of the evaluation results or add information to rectify any factual errors. It is important, however, that the evaluator(s) maintain their independence at all times during the discussions and be prepared to decide what modifications to introduce to the draft report. Any dissenting views should be properly recorded in the report. In the case of evaluator(s) residing outside the country where the evaluation is being conducted, the main conclusions and recommendations should be finalized before the evaluator(s) leave the country. The evaluation report should normally be finalized two weeks after the conclusion of the evaluation exercise and submitted to the country office.

Annex 1. Evaluation Report: Suggested Outline

Title page

- Name of project, programme or theme being evaluated.
- Country/ies of project/programme or theme.
- Name of the organization to which the report is submitted.
- Names and affiliations of the evaluators.
- Date.

Table of Contents

Acknowledgements

- Identify those who contributed to the evaluation.

List of acronyms

Executive summary

- A self-contained paper of 1-3 pages.
- Summarize essential information on the subject being evaluated, the purpose and objectives of the evaluation, methods applied and major limitations, the most important findings, conclusions and recommendations in priority order.

Introduction

- Describe the project/programme/theme being evaluated. This includes the problems that the interventions are addressing; the aims, strategies, scope and cost of the intervention; its key stakeholders and their roles in implementing the intervention.
- Summarize the evaluation purpose, objectives, and key questions. Explain the rationale for selection/non selection of evaluation criteria.
- Describe the methodology employed to conduct the evaluation and its limitations if any.
- Detail who was involved in conducting the evaluation and what were their roles.
- Describe the structure of the evaluation report.

Findings and conclusions

- State findings based on the evidence derived from the information collected. Assess the degree to which the intervention design is applying results based management principles. In providing a critical assessment of performance, analyse the linkages between inputs, activities, outputs, outcomes and if possible impact. To the extent possible measure achievement of results in quantitative and qualitative terms. Analyse factors that affected performance as well as unintended effects, both positive and negative. Discuss the relative contributions of stakeholders to achievement of results.
- Conclusions should be substantiated by the findings and be consistent with the data collected. They must relate to the evaluation objectives and provide answers to the evaluation questions. They should also include a discussion of the reasons for successes and failures, especially the constraints and enabling factors.

Lessons learned

- Based on the evaluation findings and drawing from the evaluator(s)' overall experience in other contexts if possible provide lessons learned that may be applicable in other situations as well. Include both positive and negative lessons.

Recommendations

- Formulate relevant, specific and realistic recommendations that are based on the evidence gathered, conclusions made and lessons learned. Discuss their anticipated implications. Consult key stakeholders when developing the recommendations.
- List proposals for action to be taken (short and long-term) by the person(s), unit or organization responsible for follow-up in priority order.
- Provide suggested time lines and cost estimates (where relevant) for implementation.

Annexes

- Attach ToR (for the evaluation).
- List persons interviewed, sites visited.
- List documents reviewed (reports, publications).
- Data collection instruments (e.g., copies of questionnaires, surveys, etc.).

Sources: Adapted from the UNICEF Guide for Monitoring and Evaluation, 1991; the UNICEF Evaluation Reports Standards, 2004; and the UNDP Results-Oriented Monitoring and Evaluation: A Handbook for Programme Managers, 1997.

Sources

CIDA. "**CIDA Evaluation Guide**", Performance Review Branch, 2000.

OECD. "**Improving Evaluation Practices: Best Practice Guidelines for Evaluation and Background Paper**", 1999. Available in English and French at <http://www.oecd.org/puma/budget/budpubs.htm>

Save the Children. "**Toolkits: A Practical Guide to Assessment, Monitoring, review and Evaluation**", Development Manual 5, London, 1999.

UNDP. "**Results-Oriented Monitoring and Evaluation: A Handbook for Programme Managers**", Office of Evaluation and Strategic Planning, New York, 1997.

UNICEF. "**A UNICEF Guide for Monitoring and Evaluation: Making a Difference?**", Evaluation Office, New York, 1991.

UNICEF. "**Evaluation Reports Standards**", 2004.

UNICEF. "**What goes into a Terms of Reference (ToR)**", UNICEF Technical Notes, Issue Number 2, April 2002. Available at http://www.unicef.org/evaluation/TechNote2_TOR.pdf

UNICEF. "**Writing a good Executive Summary**", UNICEF Technical Notes, Issue Number 3, August 2002. Available at http://www.unicef.org/evaluation/TechNote3_Exec_Sum.pdf

USAID. "**Performance Monitoring and Evaluation - TIPS # 3: Preparing an Evaluation Scope of Work**", 1996 and "TIPS # 11: The Role of Evaluation in USAID", 1997, Centre for Development Information and Evaluation. Available at http://www.dec.org/usaid_eval/#004

U.S. Centres for Disease Control and Prevention (CDC). "**Framework for Program Evaluation in Public Health**", 1999. Available in English at <http://www.cdc.gov/eval/over.htm>

U.S. Department of Health and Human Services. Administration on Children, Youth, and Families (ACYF), "**The Program Manager's Guide to Evaluation**", 1997.

World Food Programme. "**Policy Issues: WFP Principles and Methods of Monitoring and Evaluation**", Executive Board, Annual Session, Rome, 22-26 May 2000.

August 2004¹

PART V: COMMUNICATING AND USING EVALUATION RESULTS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This part V of tool number 5 suggests steps and considerations for the effective communication and use of evaluation results. The content is based on a review of literature both from academia and international development agencies such as UNICEF, UNDP and bilateral donor agencies such as DANIDA, OECD, USAID and the U.S. Centers for Disease Control.

II. Assessing the quality of evaluations

One of the first tasks after the evaluation is completed is to disseminate its results to potential users. It is essential, however, to have already ascertained that the evaluation has produced timely and credible information and well-founded recommendations (see Tool number 5, Part VI on evaluation Standards). Poor evaluations should not be used, but only after careful consideration of what went wrong.

III. Sharing evaluation results

It is not sufficient to merely conduct evaluation as an information-gathering activity. It is also important to disseminate and communicate evaluation results to key stakeholders and other audiences as soon as possible and in forms that are easy to understand and use. First, disseminate results to those with a direct interest in the programme being evaluated - especially to those with responsibility for making important decisions about the programme - and then to other potential users. In addition to delivering a final report, it is useful to organize meetings with various stakeholders using a variety of techniques such as visual displays and oral presentations to communicate evaluation results, and help users assimilate and interpret the information. For instance, the presenter (e.g., either the evaluator or the evaluation manager) can prepare information on the most important evaluation findings, and describe what the findings imply for programme implementation, redirection, funding, and expansion, as appropriate. Seminars, workshops and discussion groups can also be organized. Such working sessions offer opportunities for stakeholders not only to hear about evaluation findings, but also interpret them and construct meaning - these are opportunities for learning. In some cases, it may be worthwhile publishing the evaluation report either in its entirety or in shorter versions for dissemination to a wider audience (see **Box 1**. Dissemination channels to consider). Ensuring openness and a wider access to evaluation results increases their credibility and use.

¹ This tool was first published in May 2001.

Box 1. Dissemination channels to consider

- Detailed written report
- Executive summary, summaries of evaluation findings and key conclusions
- Brochure on the principal evaluation lessons and recommendations
- Annual report
- Article in technical or organizational newsletter
- News release
- Press conference
- Media appearance
- Public meeting, public debate
- Seminar, workshop, or group discussion
- Electronically (e-mail, Internet, websites)

Using different dissemination channels is vital to enhancing organizational learning and sharing of experiences across organizations and with broader audiences.

IV. Follow-up

Disseminating evaluation results does not ensure implementation of recommendations and use of lessons learned. Active follow-up is necessary to implement recommendations made to programme managers, and incorporate lessons learned in future decision-making processes such as the development of a new programme. At the conclusion of an evaluation, it is good practice for the evaluation manager to organize a meeting with the appropriate persons and institutions to establish an implementation plan based on the recommendations with a timetable and identification of parties responsible for follow-up actions (See **Box 2**. Using evaluation results). The more stakeholders are involved in planning the next steps, the more likely they are to follow through on implementing evaluation recommendations. The evaluation manager should monitor the status of implementation and by so doing advocate the use of evaluation results.

Box 2. Using evaluation results

It is the responsibility of programme managers to:

- Systematically review the key evaluation findings, conclusions and recommendations
- Identify which are accepted and supported and which are not
- In the case of a formative evaluation, determine whether any adjustments are necessary (i.e., in the programme strategy, the results and resources framework, or activities) to improve programme effectiveness
- Establish an implementation plan, including the identification of specific follow-up actions and assignment of clear responsibilities with a timetable for undertaking them
- Monitor the status of implementation.\

Source: Adapted from USAID, Performance Monitoring and Evaluation, TIPS # 11, 1997.

Sources

ILO. "**Guidelines for the Preparation of Independent Evaluations of ILO Programmes and Projects**", Evaluation Unit, 1997. Available at

<http://www.ilo.org/public/english/bureau/program/guides/evalmenu.htm>

Ministry of Foreign Affairs, DANIDA. "**Evaluation Guidelines**", February 1999.

OECD. "**Improving Evaluation Practices: Best Practice Guidelines for Evaluation and Background Paper**", 1999.

Preskill, Hallie and Torres, Rosalie T. "**The Learning Dimension of Evaluation Use**", New Directions for Evaluation, American Evaluation Association, Number 88, Winter 2000.

Rossi, Michael H., Freeman, Howard E., and Lipsey, Mark W. "**Evaluation - A Systematic Approach**", 6th edition, Sage Publications, California, 1999.

UNDP. "**Results-Oriented Monitoring and Evaluation: A Handbook for Programme Managers**", Office of Evaluation and Strategic Planning, New York, 1997.

UNICEF. "**A UNICEF Guide for Monitoring and Evaluation: Making a Difference?**", Evaluation Office, New York, 1991.

USAID. "**Performance Monitoring and Evaluation - TIPS # 11: The Role of Evaluation in USAID**", 1997, Centre for Development Information and Evaluation.

U.S. Centres for Disease Control and Prevention (CDC). "**Framework for Program Evaluation in Public Health**", 1999. Available at <http://www.cdc.gov/eval/over.htm>

U.S. Department of Health and Human Services. Administration on Children, Youth, and Families (ACYF), "**The Program Manager's Guide to evaluation**", 1997.

August 2004¹

PART VI: EVALUATION STANDARDS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This part VI of the tool number 5 lists evaluation standards to be applied throughout the evaluation process to ensure the quality of the evaluation product. In addition to a review of the literature from bilateral and development agencies such as OECD, DANIDA and U.S. Centers for Disease Control and Prevention, the content is based on Program Evaluation Standards developed by the Joint Committee on Standards for Educational Development (1994, 1999) and the adapted evaluation standards recommended by the Swiss Evaluation Society (SEVAL).

II. Ensuring the quality of evaluations

For evaluations to be useful they should meet certain standards of quality. The international community of evaluators has established standards for sound and fair evaluation, which can be applied while planning an evaluation and throughout its implementation. Some of these standards are considered as universal while others are perhaps more unique to certain cultural settings. As such, their application should be adapted taking into account the specific situation. These standards are organized around four important attributes of evaluation:

Utility Standards

The Utility standards should ensure that an evaluation is guided by the information needs of its users. These standards are as follows:

Stakeholder Identification - Persons involved in or affected by the evaluation should be identified so that their interests and needs can be addressed. The following persons, groups, and institutions are referred to as "stakeholders" and should be consulted in the context of an evaluation:

- Those who decide upon the future of the programme (often the funding agency)
- Those who are responsible for the planning and design of the programme
- Those who are involved in the implementation of the programme
- Those who should or will be directly or indirectly affected by the programme (target groups and their social contexts)
- Other groups with an interest in the evaluation findings (e.g., decision makers who plan similar programmes, evaluators, and the general public).

¹ This tool was first published in May 2001.

Evaluator Credibility - Those conducting an evaluation should be both trustworthy and competent, so that the evaluation findings achieve maximum credibility and acceptance. To be found credible by different stakeholder groups, the following characteristics are crucial: professional competence, integrity, independence, as well as social and communication skills.

Information Selection - The information collected should be comprehensive enough to address pertinent questions about the programme and be responsive to the interests and needs of stakeholders. When planning an evaluation, it is also important to distinguish information that is essential versus information that is desirable.

Transparency of Assessment - The perspectives, rationale, and procedures used to interpret the findings should be carefully described, so that the bases for value judgments are clear.

Report Clarity - Evaluation reports should clearly describe the programme being evaluated including its context, and the purposes, questions, procedures, and findings of the evaluation. The language should be precise (e.g., clear definitions of the most important terms and consistent use of terminology) and easily understood by the intended audience.

Report Timeliness - Significant interim findings and final reports should be brought to the attention of intended users, so that they can be used in a timely fashion. Evaluations are most useful when planned to fit into the stakeholders' decision making processes. For many evaluations it is sensible to share interim findings with the stakeholders, especially when these results might have an impact on their future actions.

Evaluation Impact - Evaluations should be planned, conducted, and reported in ways that encourage stakeholder participation to varying degrees, so that the likelihood that evaluation results will be used is increased. The more involved stakeholders are at the different stages of the evaluation process, the greater the likelihood they will act on the evaluation recommendations.

Feasibility Standards

The Feasibility standards should ensure that an evaluation is carried out in a realistic, thoughtful, tactful, and cost effective manner. These standards are as follows:

Practical Procedures - Evaluation methods and instruments should be practical to keep disruption to a minimum while the needed information is collected. It is crucial to discuss with stakeholders the advantages and disadvantages of chosen methods.

Political Viability - The evaluation should be planned and conducted taking into account the different positions of the various interest groups, in order to obtain a balanced presentation of different points of view. It should enlist their cooperation and avert or counteract possible attempts to curtail evaluation activities or to bias the results.

Cost Effectiveness - Evaluations should produce information of sufficient value for informed decision-making, learning and accountability so that the resources expended can be justified.

Propriety Standards

The Propriety standards should ensure that an evaluation is conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation as well as those affected by its results. These standards are as follows:

Formal Agreement - Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that they are obligated to adhere to all conditions of the agreement or to renegotiate it. Such a formal written agreement should at least regulate budget, time, personnel, design, methodology and report contents.

Protection of Individual Rights - Evaluations should be designed and conducted in a way that respects and protects the rights and welfare of human beings. If an evaluation leads to well founded conclusions that pose a threat to the welfare of individuals, the extent to which these findings are disseminated needs to be carefully considered and justified.

Human Interactions - Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation so that participants are not threatened or harmed. This is not only a matter of human dignity but also relates to practical considerations. It is therefore necessary to be familiar with the cultural practices (i.e. beliefs, manners and customs) of those involved.

Complete and Fair Assessment - Evaluations should be complete and fair in their examination and recording of strengths and weaknesses of the programme being evaluated, so that strengths can be built upon and problem areas addressed. If, for whatever reason (e.g., because of time or budgetary constraints), there are issues that cause difficulties for the evaluation (e.g., it was impossible to collect certain data) these should be brought to light.

Disclosure of Findings - The formal parties to an evaluation should ensure that the full set of evaluation findings is made accessible to the persons affected by and/or interested in the evaluation.

Conflict of Interest - Conflict of interest should be dealt with openly and honestly so that it does not compromise the evaluation process and results. It is therefore crucial that evaluators be able to clarify their roles and make a distinction between facts and opinions. The integrity of the evaluation cannot be compromised just to accommodate conflicts of interest.

Accuracy Standards

The Accuracy standards should ensure that an evaluation would reveal and convey technically adequate information about the features that determine the value of the programme being evaluated. These standards are as follows:

Programme Documentation - The Programme being evaluated should be described and documented clearly and accurately. The description should be sufficiently detailed to ensure an understand-

ing of programme aims and strategies. It is especially crucial to note differences between the planned and the actual performance of the programme.

Context Analysis - The context in which the programme exists should be examined in enough detail so that its likely influences on the programme can be identified. Understanding the setting in which a programme functions will help in the accurate interpretation of evaluation findings and in assessing the extent to which they can be generalized.

Described Purposes and Procedures - The purposes and procedures of an evaluation should be monitored and described in enough detail so that they can be identified and assessed. It is important that the evaluation process focus on the issues of greatest concern to stakeholders while using time and resources as efficiently as possible.

Defensible Information Sources - The sources of information used in a programme evaluation should be described in enough detail so that their adequacy can be assessed. The criteria used for selecting sources should be stated clearly so that users and other stakeholders can interpret the information accurately and assess if it might be biased.

Valid and Reliable Information - The information gathering procedures implemented should provide assurance that the interpretation arrived at is valid and reliable. Validity is defined by the extent to which methodologies and instruments measure what they are intended to measure. A data collection method is reliable to the extent that it produces the same results repeatedly.

Systematic Review of Information - The information collected, analyzed, and reported in an evaluation should be systematically reviewed and any errors found should be corrected.

Analysis of Qualitative and Quantitative Data - Qualitative and quantitative data should be analyzed in an appropriate, systematic way so that the evaluation questions can be effectively answered. Data analysis should follow rules of methodological soundness.

Justified Conclusions - The conclusions reached in an evaluation should be explicitly justified so that stakeholders can assess them. Evaluation conclusions are justified when they are based on a synthesis of empirical findings derived from the information collected. Evaluation information must be interpreted to appreciate the practical significance of what has been learned. Conclusions can be both positive and negative. Controversial conclusions should be substantiated.

Impartial reporting - Reporting procedures should guard against distortion caused by personal feelings and biases of any stakeholder group. All relevant perspectives need to be fairly represented.

Meta evaluation - The evaluation itself should be subject to an assessment of the evaluation's process and quality upon its completion using these and other pertinent standards to determine its strengths and weaknesses.

Sources

American Evaluation Association. **"The Programme Evaluation Standards"**. Available at: <http://www.eval.org/EvaluationDocuments/progeval.html>

Centres for Disease Control and Prevention (CDC). **"Framework for Program Evaluation"**, 1999.

CIDA. **"CIDA Evaluation Guide"**, Performance Review Branch, 2000. Available at <http://www.acdi-cida.gc.ca/index-e.htm>

DANIDA. **"Evaluation Guidelines"**, Ministry of Foreign Affairs, 1999.

OECD. **"Improving Evaluation Practices: Best Practice Guidelines for Evaluation and Background Paper"**, 1999.

Rossi, Michael H., Freeman, Howard E., and Lipsey, Mark W. **"Evaluation - A Systematic Approach"**, 6th edition, Sage Publications, California, 1999.

Swiss Evaluation Society (SEVAL). **"Evaluation Standards"**, December 2000. Available at: <http://www.seval.ch/en/standards/index.cfm>

UNDP. **"Results-Oriented Monitoring and Evaluation: A Handbook for Programme Managers"**, Office of Evaluation and Strategic Planning, New York, 1997.

UNICEF. **"A UNICEF Guide for Monitoring and Evaluation: Making a Difference?"**, Evaluation Office, New York, 1991.

USAID. **"Performance Monitoring and Evaluation - TIPS # 3: Preparing an Evaluation Scope of Work, and TIPS # 11: the Role of Evaluation in USAID"**, 1997, Centre for Development Information and Evaluation.

Western Michigan University. **"The Program Evaluation Standards in International Settings"**, The Evaluation Center, Occasional Papers Series, May, 2000.



6

Tool Number

PROGRAMME INDICATORS



Division for Oversight Services

August 2004¹

PROGRAMME INDICATORS

PART I: IDENTIFYING OUTPUT INDICATORS

THE BASIC CONCEPTS

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results-based management. It is also useful for other programme managers at headquarters and for national programme managers and counterparts.

Tool number 6 was produced in collaboration with the UNFPA Technical Support Division. It provides guidance to UNFPA and other programme managers on the basic principles in identifying sound output indicators and Means of Verification (MOVs)².

With UNFPA's adoption of results-based management as a guiding principle for programming, indicators have become important instruments for UNFPA to measure the results of its development assistance at the output, outcomes and goal levels. Indicators, particularly those at the output level of results, for which UNFPA's programmes are responsible, provide essential information for successful programme management. In-house reviews of country programme logframes have found that the Logical Framework/ Results and Resources Framework indicators at the output level and their associated MOVs are often not well identified³.

II. The Process

In order to ensure that important programme stakeholders use UNFPA funded programme output indicators to track programme results, it is essential that they participate in the indicator selection process. The process should be initiated as part of the definition of programme results carried out together with main programme stakeholders. It is important to note that the establishment of an adequate set of indicators to track programme results is an iterative process whereby the set of indicators and performance targets is improved and adjusted, particularly in the early years of programme implementation, as the availability of baseline data improves. It should be emphasized that indicators have to be practical and related to this, that steps are taken to ensure that systems for collecting the necessary data (means of verification) are in place and are funded.

¹ This tool was first published in August 2002.

² For a discussion and listing of purpose and goal level indicators, please consult the UNFPA publication, *Indicators for Population and Reproductive Health (outcome) Programmes*, (1998).

³ Box 3, page 5, provides a review of commonly encountered problems.

Step 1: Define the planned outputs

Output indicators should tell us how the programme is performing. They are the detailed expressions of the programme results for which UNFPA is responsible. Thus, before programme stakeholders identify output indicators, they must reach a consensus on the content of the output statements. **Box 1** shows how to.

Box 1. How to clarify the outputs?

- Review the wording and intention of the output. What exactly does it say?
- Avoid broad output statements. They make indicator identification difficult.
- Be clear about what type of change is implied in the output. What is expected to change - a condition, level of knowledge, attitude?
- Be clear about where change should appear. Among individuals, counterpart organizations, communities, regions?
- Specify in more detail the targets for change. Who or what are the specific targets for change? What kind of individuals, organizations etc.?

Source: Adapted from USAID TIPS 1996, number 6.

The following examples taken from UNFPA logframe matrices/results and resources framework illustrate typical output statements:

"improved access to/availability of quality RH services"

"strengthened capacity of MOH to plan and manage RH services"

"improved awareness and knowledge on reproductive rights and reproductive health of women, men and adolescents"

These statements lack clarity and specificity; they leave considerable room for interpretation by programme implementers and managers:

What is meant by "access"? "quality" RH services? What elements of RH services are being targeted? What is "strengthened planning and management capacity"? What is the difference between "awareness" and "knowledge"? What is "reproductive rights"? Are these outputs to be achieved for the whole country or special target areas? The public or the private sector? Any target population?

To identify appropriate indicators to track these outputs we need to know:

- The target area
- The target population⁴
- The specific RH services⁵

⁴ Women, men and adolescents and sub-groups of these sharing common characteristics, socio-economic groups.

- The kind of access⁶
- The quality standards for each specific RH service component⁷
- The specific planning and management aspects which need to be strengthened⁸
- The specific awareness or knowledge to be developed.

Some output elements, such as target groups or geographical location, can be specified very briefly in a footnote to the logframe/results and resources framework or in the indicators. Other output elements, such as quality of care or specific knowledge and attitudes expected from the target groups, are more complex and may require more work after elaboration of the logframe/results and resources framework for example to identify specific standards and to incorporate them in checklists that can be used for monitoring indicator progress in the course of programme implementation.

Step 2: Identify the best indicator or cluster of indicators and the performance targets to track each output

How to identify good indicators?

In general good indicators need to be:

- relevant to the programme
- relevant to national standards
- feasible to collect
- easy to interpret
- should enable tracking of change over time.

Various organizations use different criteria to select relevant indicators⁹. DOPA criteria (explained in Box 2) encapsulate the most important requirements of useful indicators. They are a simple tool to guide

⁵ Essential obstetric care; family planning; prevention and management of RTI/STD and prevention of HIV infection; management of the consequences and complication of unsafe abortion; information, education and counselling.

⁶ Geographic/physical distance; economic; administrative; cognitive; psychosocial and cultural.

⁷ Provider performance/service delivery according to protocols; performance of systems such as training, supervision, management, health information system, logistics; client satisfaction.

⁸ Such as the presence of a strategic plan; of a system for preparing yearly operational plans for the organization; of a regular system to assess the needs and preferences of clients and to adjust services in response to identified changes; of a manager whose job description includes assessing clients' needs, developing the strategic and operational plan, revising and assessing the operationalisation of the plan. For more details see USAID: Health and Family Planning Indicators: Measuring Sustainability, Volume II, available at http://sara.aed.org/publications/cross_cutting/indicators/html/indicators.htm

⁹ Many organizations use SMART which stands for: Specific (measures exactly the result); Measurable (so that the result can be tracked); Attainable (realistic); Relevant (to the intended result) and Timebound (indicates a specific time period).

Box 2. What are DOPA Criteria?

They are standards used to assess that the indicators are:

Direct

- closely measure the intended change.

Objective

- unambiguous about what is being measured and which data to be collected.
- clear operational definition that is independent of the person conducting the measurement.

Practical

- reasonable in terms of data collection cost, frequency, and timeliness for decision-making purposes.

Adequate

- the minimum number of indicators necessary to ensure that progress towards the output is sufficiently captured.

Source: USAID TIPS Number 12, 1998.

Box 3 illustrates common problems with respect to UNFPA funded country programme logframe/results and resources framework output indicators. Problems 1 and 6 highlight that output indicators do not correspond to the output level: they are not *direct*.

If it is not possible to use a direct measure, one or more proxy indicators might be appropriate. A proxy indicator is an indirect measure that is linked to the result by one or more assumptions. For example, the contraceptive prevalence rate can be considered a proxy indicator for access to and utilization of RH services.

Problem 2 highlights that output indicators are often ambiguous or lacking detail: they are not *objective*.

Therefore there may not be consensus among programme and project implementers and managers about what is being measured and what are the results of an intervention.

Problems 4 and 5 highlight that UNFPA logframe matrices/results and resources framework often include too many and/or unrealistic output indicators, which it would be impossible to adequately manage and which are not all necessary to capture progress in achieving the output: they are neither *practical* nor *adequate*. Instead, programme stakeholders should select one or a cluster of two or three indicators which are *practical* because they allow managers to track the output in the most direct and timely fashion and require the least effort in terms of time, human and financial resources, and *adequate* because together they represent the most important dimensions of the planned output.

Box 3. Common Problems in Specifying Output Indicators

A desk review of UNFPA Country Programme documents highlighted some weaknesses in specifying output indicators:

1. Indicators **do not correspond to the output level**. For example, for an output of "improved availability of RH services" use of an **activity indicator**, such as quantity of equipment procured, or of an **outcome indicator**, such as the contraceptive prevalence rate. Quantity of equipment procured does not alone ensure that RH services are available. Additionally, availability of RH services does not necessarily lead to increased use of such services and therefore increased CPR.
2. Indicators **do not include an objective standard** against which achievement can be assessed. For example, an indicator of "system developed" is used for an output such as "system for coordination, monitoring and evaluation of population programmes". The standard needs to be defined explicitly.
3. Indicator targets **without reference to a baseline**.
4. **Too many indicators** with little consideration of the time, human resources and cost required to collect the indicator data.
5. Indicators that seem **unrealistic** due to lack of data to construct the specified indicator and/or because the indicator is very difficult to measure.
6. **Inconsistency between the universe of the output and the indicators**. For example, the **output** relating to a few sample areas but the specified indicators relating to the entire country.
7. **Copying of indicators contained in UNFPA guidelines** without consideration of their relevance to the specific programme context.
8. **Infrequent** use of **gender sensitive indicators**.

Box 4 shows how to narrow down the selection of indicators for specific programme outputs.

Box 4. Identifying an adequate and practical cluster of indicators

1. Identify a limited number of indicators which adequately measure the outputs.
2. Identify the data sources available and the type of data collection needed for each indicator.
3. Construct a matrix listing the indicators, identifying their importance for programme monitoring (high/low), the ease of obtaining data on the indicator (easy/feasible but requires effort/difficult), and the cost of data collection.
4. Prioritize indicators by importance, ease of obtaining data and cost and select a set of indicators.
5. Group selected indicators by source of data to determine the set of sources which can provide data on clusters of indicators.
6. Make a final selection of a cluster of indicators and decide on a data collection plan in light of available logistical, human and financial resources and time.

Source: Adapted from Bertrand and Tsui, 1995.

What is the difference between an indicator with and without a target?

Indicators tell us what we are measuring; targets are the results expected in the context of the specific programme and within a certain time frame (see example in **Box 5**). While UNFPA's guidelines require that the programme logframe/results and resources framework output indicators include targets, this has often not been possible for lack of sufficient knowledge on the baseline situation at the time of preparing the programme logframe matrix/results and resources framework.

Thus, the output indicators currently stated in most of UNFPA's logframe matrices/results and resources framework do not include targets. It is expected that it will be easier to identify targets in the course of programme implementation and in the development of future country programmes as the required surveys and/or research will then have been completed.

Box 5. Examples of an Indicator with and without a Target

Indicator without a Target

Number of service delivery points (SDPs) per population of reproductive age in each priority district where a package of minimum three types of clinical services and related IEC and counselling activities are offered.

Indicator incorporating a Target

500 SDPs/1.5 million population of reproductive age in the three district of (names) offer FP, Maternal Health and STI preventive and curative services as well as related interpersonal counselling, group communication activities and information materials by 2006.

What are different types of indicators and targets?

Indicators and targets may express quantity (how much), quality (how good), or efficiency (best output at lowest cost). **Box 6** illustrates common ways of expressing these different types of indicators and targets. Each type of indicator and target conveys a different dimension of the planned output. For example, quantitative indicators and targets provide "hard data" to demonstrate results achieved. They also facilitate comparisons and analysis of trends over time. Qualitative indicators and targets provide insights into changes in organizational processes, attitudes, beliefs, motives and behaviours of individuals¹⁰.

Qualitative indicators and targets must be expressed quantitatively in order to illustrate change. This can for instance be done by using a scoring system. A scoring system to track improvement in the quality of RH services over time could include an indicator such as "the percent of existing SDPs with a score of four out of a total of five points on a quality of care checklist increased from X to Y".

Efficiency indicators should tell us if we are getting the best value for our investment. In order to establish such an indicator, we need to know the "market", i.e. the current price of desired outputs considering both quantity and quality aspects.

How can we identify targets?

In setting targets it is important to be realistic about the outputs that are feasible to achieve given contextual constraints and past experiences in a particular sector. **Box 7** provides a few suggestions of useful information for target setting.

Box 6. Common ways of expressing Quantitative, Qualitative and Efficiency Indicators/Targets

Quantitative indicators/targets are statistical measures

- Number
- Percent
- Rate (ex. birth rate - Births per 1,000 population)
- Ratio (ex. sex ratio - Number of males per number of females).

Qualitative indicators/targets imply qualitative assessments

- Compliance with
- Quality of
- Extent of
- Level of

Efficiency indicators/targets are unit cost measures

- Cost per unit of (clients treated, student, school etc.).

¹⁰For a related review of qualitative and quantitative data collection methods, see the Programme Manager's Planning, Monitoring and Evaluation Toolkit Number 5, Part III: Planning and Managing and Evaluation - The Data Collection Process at the UNFPA's website www.unfpa.org

Box 7. Useful information for establishing targets

- Baseline data indicating the situation at the beginning of programme implementation. When such data is not available, the programme should include an activity to collect it from the start.
- Historical trends in the indicator value over time. What pattern of change has been evident in the past? Is this pattern likely to continue?
- Stakeholders' expectations of progress. Exploring the achievement expectations of national counterparts such as programme implementers and managers may be useful to provide a realistic idea of what can be achieved.
- Expert judgements and research findings. Experts knowledgeable about the programme sector and local conditions as well as research findings are other useful sources of information for target setting.
- Accomplishments of similar programmes. Information on what is being done in the programme sector under similar conditions by other agencies and organizations who have a reputation for high performance is an excellent input to the target setting process.

Source: USAID TIPS Number 8, 1996.

Step 3: Identify the Means of Verification (MOVs), timing and reporting responsibility

The indicator MOV is the data that is needed to determine the value of the indicator. MOV data can be collected through review of documentation, facility observation, in-depth interviews, focus group discussions, small surveys such as facility based quality of care surveys. In order to save time and effort, the feasibility of using existing data systems and sources, such as Demographic and Health Surveys, Health Information Systems, government or NGO administrative records, to provide some of the output indicator data should be explored. Essential output data for which there are no existing sources should be collected as an activity, for instance rapid appraisal surveys, in the context of the UNFPA funded programme. It is important, before planning a data collection activity in the UNFPA programme, to check the data collection plans of other actors or partners in the concerned sector. For instance, UNFPA may negotiate the collection of data useful for tracking programme progress as part of other UN agencies' data collection efforts such as the UNICEF sponsored Multi-cluster indicators surveys, thereby saving time and money.

All UNFPA funded programmes must indicate the MOVs as well as the timing of data collection and organizational units responsible for aggregating, analyzing, and using the data to report on progress in relation to the planned outputs.

In order to efficiently track progress in programme achievements over time and to enhance institutional memory as well as enable sharing of data among stakeholders, it is extremely useful to establish a programme database. Such a database can be part of a national population programme database, or if such does not exist, could be established within the UNFPA Country Office or the main UNFPA programme partner¹¹.

¹¹ A description of programme databases established by UNFPA Cameroon and Senegal is available in UNFPA: Setting up an Integrated Programme Database (IPDP): a shared experience of two country offices: Senegal and Cameroon, 2004.

Time, money and responsibilities must be allocated in all UNFPA funded projects for these important indicator and data management activities. The process of detailing responsibilities and a budget for data collection and management activities will often lead to a re-examination of data priorities.

III. Conclusion

Box 8 summarizes good practices in identifying and managing indicators that have been discussed in this tool. Some of these practices also apply to efficient management of outcome and impact indicators.

Box 8. Good Practices in Identifying Indicators

- **Ownership.** Involve key stakeholders in the selection of the indicators that will be used to measure programme performance;
- **Start with programme design.** Implications for data collection need to be fully integrated in the design of the programme, including a budget to cover data collection costs;
- **Baseline information.** Where change is being assessed obtain baseline data at the start of the programmes, and, if possible, data on past trends;
- **Use existing data sources and reporting systems where possible.** However, if data is not available, cost-effective and rapid assessment methodologies should be considered for supplementary data collection;
- **Partnerships.** Establish partnerships with government, NGOs, bilateral donors and other key stakeholders to collect the data so as to reduce costs;
- **Information management.** Plan how the flow of information relating to the indicators will be managed, stored and retrieved in a user-friendly data base within the Country Office or in the main UNFPA counterpart organization.

Sources

Bertrand, Jane and Amy Tsui. **"Indicators for RH Program Evaluation"**, The Evaluation Project. University of North Carolina at Chapel Hill, 1995.

Website: <http://www.measureprogram.org/>

USAID. **"Health and Family Planning Indicators: Measuring Sustainability"**, Volume II, Africa Bureau's Office of Sustainable Development (AFR/SD). Available online in English at http://sara.aed.org/publications/cross_cutting/indicators/html/indicators.htm

UNFPA. **"Measuring Outputs in Population and Development with Illustrative Indicators"**, Technical Support Division, November 2001.

USAID. Performance Monitoring and Evaluation TIPS, Centre for Development Information and Evaluation. Available online in English at <http://www.dec.org/usaideval/#004>

TIPS Number 6, 1996: Selecting Performance Indicators.

TIPS Number 8, 1996: Establishing Performance Targets.

TIPS Number 12, 1998: Guidelines for Indicator and Data Quality.

UNFPA. Training materials from Programme Management Workshops accessible in English, French and Spanish at <http://bbs.unfpa.org/personnel/training.htm>

List of Indicator Resources

I. Indicator Tools

** Adamchak, Susan; Bond, Katherine; MacLaren, Laurel; Magnani, Robert; Nelson, Kristin; Seltzer, Judith. **"A Guide to Monitoring and Evaluating Adolescent Reproductive Health Programs"**, Focus Tool Series 5, Pathfinder, June 2000.

Website: http://pf.convio.com/site/PageServer?pagename=Publications_FOCUS_Guides_and_Tools

Canadian International Development Agency (CIDA). **"Guide to Gender-sensitive Indicators"**, August 1997. Website: <http://www.acdi-cida.gc.ca/equality>

** IPPF, Western Hemisphere Region. **"Manual to Evaluate Quality of Care from a Gender Perspective"**, January 2000.

Website: http://www.ippfwhr.org/publications/publication_detail_e.asp?PubID=7

** Maine, Deborah et al. **"Guidelines for Monitoring the Availability and Use of Obstetric Services"**, UNICEF, WHO, UNFPA, August 1997.

Website: <http://www.unicef.org/health/files/guidelinesformonitoringavailabilityofemoc.pdf>

** MEASURE Evaluation. "**Resource on Population, Health and Nutrition Indicators**".

Website: <http://www.measureprogram.org/>

Morgan, Peter. "**The Design and Use of Capacity Development Indicators**". Paper prepared for the Policy Branch of CIDA, December, 1997.

** Nelson, Kristin; MacLaren, Laurel; Magnani, Robert. "**Assessing and Planning for Youth-Friendly RH Services**", Focus Tool Series, Pathfinder, January 2000.

Website: <http://pf.convio.com/site/PageServer?pagename=Publications FOCUS Guides and Tools>

Shah, Meera Kaul; Zambezi, Rose; Simasiku, Mary. "**Listening to Young Voices: Facilitating Participatory Appraisals on Reproductive Health with Adolescents**", Pathfinder.

Website: <http://pf.convio.com/site/PageServer?pagename=Publications FOCUS Guides and Tools>

** UNAIDS. "**National AIDS Programmes, A Guide to Monitoring and Evaluation**", June 2000. Website: <http://www.cpc.unc.edu/measure>

** USAID. "**Health and Family Planning Indicators: A Tool for Results Frameworks**", Volume I, Africa Bureau's Office of Sustainable Development (AFR/SD), July 1999.

Website: www.usaid.gov/regions/afr/hhrra/indicators/indicators1.htm

** USAID. "**Health and Family Planning Indicators: Measuring Sustainability**", Volume II, Africa Bureau's Office of Sustainable Development (AFR/SD).

Website: http://sara.aed.org/publications/cross_cutting/indicators/html/indicators.htm

II. Indicator Manuals/Guidance Notes

Danida. "**Output and Outcome Indicators System Manual**", November 1999.

Morgan, Peter. "**An Update on the Performance Monitoring of Capacity Development Programs - What are we Learning?**". Paper presented at the meeting of the DAC Informal Network on Institutional and Capacity Development held in Ottawa, May 3-5th, 1999.

UNDP. "**Selecting Key Results Indicators, Suggestions and Practical Advice from EO and OSG**", May 1999.

Website: http://stone.undp.org/undpweb/eo/evalnet/docstore3/yellowbook/documents/key_indicators.pdf

** USAID. "**Establishing Performance Targets**", Performance Monitoring and Evaluation TIPS Number 8, 1996. Website: http://www.dec.org/pdf_docs/pnaby226.pdf

** USAID. "**Guidelines for Indicator and Data Quality**", Performance Monitoring and Evaluation TIPS Number 12, 1998. Website: http://www.dec.org/pdf_docs/pnac927.pdf

** USAID. "**Measuring Institutional Capacity**", Recent Practices in Monitoring and Evaluation TIPS Number 15, 2000. Website: http://www.dec.org/usaid_eval/#004

** USAID. "**Selecting Performance Indicators**", Performance Monitoring and Evaluation TIPS Number 6, 1996. Website: http://www.dec.org/pdf_docs/pnaby214.pdf

World Bank. "**Performance Monitoring Indicators - A handbook for task managers**", Operations Policy Department, Washington D.C. 1996

** Indicates recommended resources

August 2004¹

PART II: INDICATORS FOR REDUCING MATERNAL MORTALITY

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA country office staff to improve planning, monitoring and evaluation activities. It is also useful for other programme managers at headquarters and national levels.

Each year more than 500,000 women die from pregnancy related causes. Most of these deaths (98%) occur in the developing world, and nearly all could be prevented. Maternal mortality reduction was highlighted at the International Conference on Population and Development (ICPD) in 1994 and at its review in 1999 (ICPD+5), and was selected as one of eight primary development goals at the United Nations Millennium Summit in 2000. The 5th Millennium Development Goal calls for a reduction of the maternal mortality ratio by three quarters between 1990 and 2015. This is an achievable goal, but only if interventions are planned based on validated strategies that successfully reduce the incidence of maternal deaths.

Tool Number 6, Part II was produced by the Technical Support Division, UNFPA, in collaboration with the Division for Oversight Services and the UNFPA Evaluation and Maternal Mortality Networks. Sections II and III of the tool highlight the major causes of maternal death and describe effective strategies to prevent its occurrence. Section IV provides indicators proven to be practical and effective both for needs assessments to plan maternal mortality interventions and for tracking progress while implementing them. It also highlights sources of data for each indicator. The indicators described in this tool are currently used in several maternal mortality reduction programmes supported by UNFPA as well as other organizations. The functionality of the indicators is being assessed through these interventions.

While the tool mentions key issues regarding demand of emergency obstetric care services such as those related to the policy environment, individual and community awareness and commitment to reducing maternal mortality, it mainly addresses issues pertaining to supply of Emergency Obstetric Care (EmOC) services. Indicators for results related to demand for maternal mortality reduction would be the subject of a future tool.

II. Medical Causes of Maternal Death

When addressing maternal mortality in any country it is essential to keep in mind several medical factors. The overwhelming majority of complications in pregnancy cannot be predicted, nor can they be prevented. Some risk factors have been determined, but it remains nearly impossible to predict which individual women will develop complications. In fact, *all* pregnant women are at random risk for developing life-threatening complications. Evidence shows that 15% of all pregnancies result in complications. Most women who develop complications have been considered to

be "low risk." The four complications that most often lead to death are haemorrhage, sepsis (infection), eclampsia (pregnancy-induced hypertension), and obstructed labour. Together with unsafe abortion, these complications are responsible for more than two-thirds of maternal deaths. The remaining third are due to indirect causes or an existing medical condition-usually malaria, anaemia, hepatitis, or AIDS-that is worsened by pregnancy or delivery.

III. Reducing the Incidence of Maternal Deaths-A Chain of Results

Learning from History

For many years, maternal mortality reduction programmes focused on two main components: antenatal care and the training of traditional birth attendants and community members to recognize complications during delivery. The intent of these programmes was that women with life-threatening complications would be transferred to a higher level of care in a timely fashion. Unfortunately, this strategy did not address several key aspects of the problem. Perhaps the most significant weakness of this approach was the failure to address the availability of care at higher levels of the system. Countries with high rates of maternal mortality nearly always have a dearth of facilities offering EmOC. Any programme that seeks to increase referrals to a higher level of care will fail if such services are not available.

Another weakness, as shown by many studies, is that traditional birth attendants simply do not have the skills to recognize complications, even when trained, because they often attend relatively few complicated deliveries during their "careers." They may also lack the authority to convince the family to transfer a labouring woman to a hospital. And, even in communities where traditional birth attendants are influential, they may still lack the skills to save women's lives in the event of a medical emergency.

Much of the current thinking about maternal mortality comes from observing countries that have been successful in dramatically reducing maternal mortality such as the United States and European countries in the 18th, 19th and 20th century¹. In Sweden, a concerted effort was made in the late nineteenth century to promote skilled attendance at birth. Midwives were trained and deployed, resulting in a precipitous drop in maternal mortality. After World War II there was a second drop in maternal mortality in both Europe and North America. This was a direct result of the introduction of antibiotics, blood transfusions and readily available Caesarean sections. These interventions are effective in preventing most causes of maternal deaths: sepsis, haemorrhage, and obstructed labour.

In several resource poor countries such as Cuba, Sri Lanka, Malaysia and Tunisia, the Western model of maternal mortality reduction has been replicated over the past 30-50 years. In these countries the establishment of national birth and death registers that include the causes of death enabled monitoring of maternal mortality trends. Analysis of improved vital statistics registers revealed the high toll of maternal deaths and brought about an awareness of the problem, which then led to an increase in political will and a swift legislative effort to improve access to skilled care at birth.

¹ Maternal mortality ratios (i.e. the number of maternal deaths per 100,000 live births) in Sweden, Great Britain and the USA reached levels of 250-400 per 100,000 live births in the mid-19th-century. A previous drop from 800-1000 had occurred between 1750 and 1850 through the gradual increase in the proportion of births attended by skilled professionals, mostly midwives.

Box 1 highlights some key factors for the success of maternal mortality interventions currently implemented in both poor and rich countries.

Box 1. Key features of successful maternal mortality reduction programmes.

- Consistent political commitment at the highest level
- Professionalisation of midwifery care
- Skilled attendance at 90% of deliveries
- Access to family planning/contraceptive services
- Access to quality Emergency Obstetric Care including medical technologies such as antibiotics, anticonvulsants and oxytocics
- Continuous service monitoring and quality improvement including maternal death audits

How and When to Intervene

Programme planners and managers are concerned with planning and implementing maternal mortality reduction programmes that work and are most likely to achieve the desired impact.

The first strategy is to prevent unwanted pregnancies from occurring at all. Then, once a pregnancy has occurred, the focus should be on the critical and dangerous time of delivery and the immediate postpartum period. Experience shows that women who die in childbirth experienced at least one of the classic three delays illustrated in **box 2**.

Political will to reduce maternal mortality is essential, since maternal mortality requires a long-term commitment of funds, infrastructure development and appropriate policies. An influx of resources to maternal health is most effective when part of an effort to strengthen the entire health sector. Indeed, in countries that experience a decline in maternal mortality the common element has been the political commitment to strengthening the health sector as a whole.

The Maternal Mortality Chain of Results (see **figure 1**) visualizes in a schematic fashion the key interventions required to reduce maternal mortality based on current knowledge of what works in maternal mortality reduction². In this context, and considering a results-based approach to programme management, the figure illustrates a change process over time and the type of results (outputs³) that UNFPA funded programmes could be responsible for delivering. The figure clarifies that donor funded programmes have increasingly reduced influence on achievement of higher level results such as outcomes and impact due to many intervening factors external to the programme. Developing a chain of results indicating the possible process through which maternal mortality could be reduced is useful to guide both programme design and progress monitoring and is an essential first step in identifying effective indicators that allow managers to track change.

² It should be noted that the chain of result depicted in figure 1 does not illustrate strategic interventions and activities necessary to achieve the outputs.

³ The following definitions have been approved by the UN Task Force on Simplification and Harmonisation of which UNFPA is a member: "Outputs are the products and services that result from the completion of activities within a development intervention. Outcomes are the intended or achieved short-term and medium-term effects of an intervention's outputs, usually requiring the collective effort of partners. Outcomes represent changes in development conditions that occur between the completion of outputs and the achievement of impact. Impacts are the positive and negative long-term effects on identifiable population groups produced by a development intervention, directly or indirectly, intended or unintended. These effects can be economic, socio-cultural, institutional, environmental, technological or health and demographic."

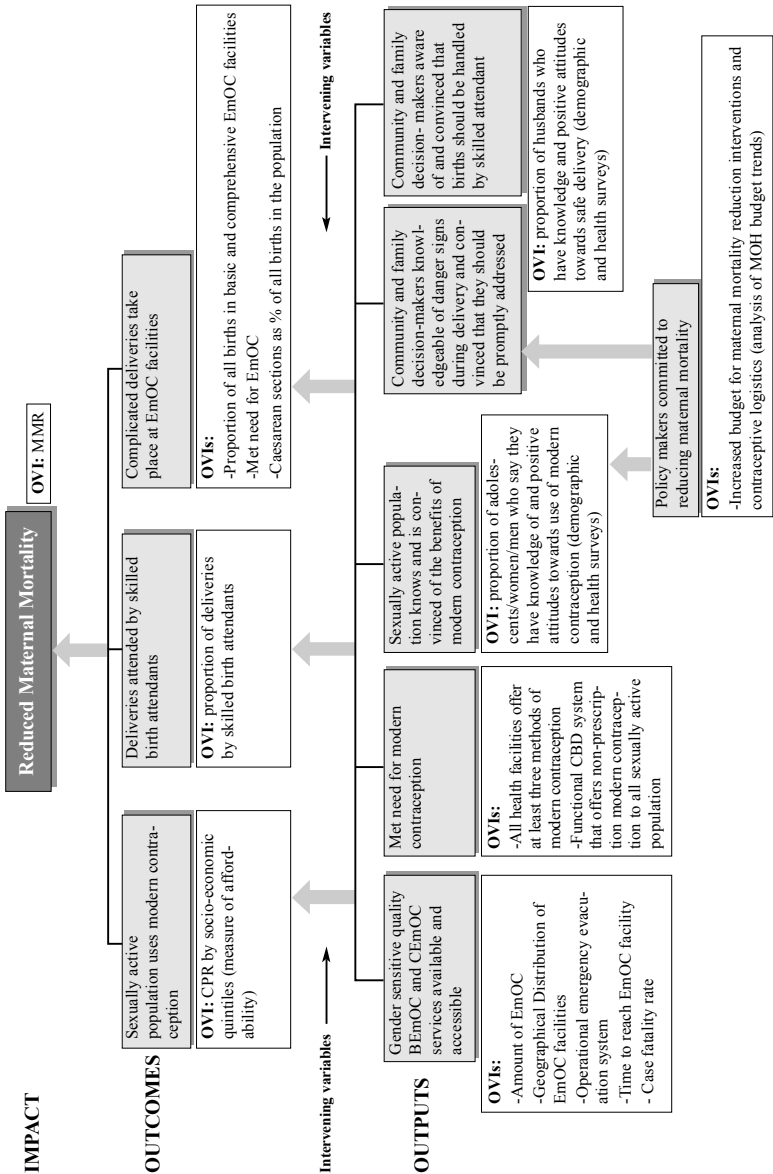
Box 2. The Three Delays

The first delay is the delay in deciding to seek care for an obstetric complication. This may occur for several reasons, including late recognition that there is a problem, fear of the hospital or of the costs that will be incurred there, or the lack of an available decision maker.

The second delay occurs after the decision has been made to seek care. This is a delay in physically reaching the care facility and is usually caused by difficulty in finding or paying for transportation. Many villages have very few transportation options and are connected by poor roads. Some communities have developed innovative ways of addressing this problem, including prepayment schemes, community transportation funds, and a strengthening of links between community practitioners and the formal health system.

The third delay is the delay in obtaining care once present at the facility. This is one of the most unfortunate issues in maternal mortality. Often, women wait for many hours at the referral centre because of poor staffing, prepayment policies, or difficulties in obtaining blood supplies, equipment or an operating theatre. The third delay is the area that many planners feel is easiest to correct. Once a woman has actually reached an EmOC facility many of the economic and socio-cultural barriers have already been overcome. Focusing on improving services in existing centres is a major component in promoting access to EmOC. Programs designed to address the first two delays (i.e. programmes that educate communities to recognise complications and encourage them to seek care, or programmes designed to improve transportation to a facility offering a higher level of care) are of no use if the facilities themselves are not adequate.

Figure 1: Chain of Results for Maternal Mortality Reduction: Results and their Objectively Verifiable Indicators (OVIs)



UNFPA's three-pronged Approach to Improving Maternal Health

Based on the knowledge of what works described above, UNFPA has adopted a three-pronged approach to reducing maternal mortality complemented by policy level advocacy and behaviour change communication interventions. The "three prongs" are described below.

1. Family Planning

Meeting the existing demand for family planning services alone would reduce pregnancies in developing countries by 20% and maternal deaths and injuries by a similar degree or more. UNFPA's strategy has been refined over the past 30 years to ensure that family planning services are of high quality; that there is an adequate supply of a wide range of contraceptives and reproductive health supplies, including male and female condoms; and, that individual choice is respected. While access to family planning will do little to reduce maternal mortality *ratios* it does a great deal to reduce the overall rate of deaths related to pregnancy and unsafe abortions⁴. FP is a cost-effective means to lower maternal mortality rates by:

1) reducing the absolute number of complications due to fewer pregnancies; 2) reducing the incidence of abortion by averting unwanted and unplanned pregnancies; 3) averting pregnancies that occur too early, too late or too frequently during the woman's reproductive cycle, and those that are inadequately spaced.

2. Skilled Attendance at Birth

Most obstetric complications occur at the time of labour and delivery. It takes a skilled attendant to swiftly recognize life-threatening complications and to intervene in time to save the mother's life. **Box 3** illustrates what is meant by "skilled attendant." In spite of overwhelming historical evidence that the use of doctors, midwives and nurses in deliveries is a crucial factor in reducing maternal mortality, only 58% of deliveries worldwide currently take place in the presence of a skilled attendant.

There are many reasons for this discrepancy. One is simply a lack of skilled attendants. Another factor is a poor distribution of attendants, with most professionals preferring to remain in urban areas. UNFPA is seeking to address this problem by promoting the training of professionals and innovative programmes to retain them in the regions of greatest need. This includes providing incentives like housing and distance learning programmes to midwives and doctors working in rural and semi-rural areas, and promoting rotation systems with a mix of public and private practice. In addition, decentralization of training that is adapted to the local context may help to retain some skilled professionals in the rural areas, especially those in the intermediate categories such as auxiliary-nurse-midwife, family welfare visitor, or lady health worker.

⁴ A **maternal mortality rate** is defined as: pregnancy related deaths per 100,000 women aged 15-49 per year. A **maternal mortality ratio** is defined as pregnancy-related deaths per 100,000 live births.

Box 3. Who is a Skilled Attendant?

The term "skilled attendant" refers exclusively to people with midwifery skills (for example, doctors, midwives, nurses) who have been trained to proficiency in the skills necessary to manage normal deliveries and diagnose or refer obstetric complications

Ideally, skilled attendants live, in, and are part of, the community they serve. They must be able to manage normal labour and delivery, recognize the onset of complications, perform essential interventions, start treatment, and supervise the referral of mother and baby for interventions that are beyond their competence or not possible in that particular setting. Depending on the setting, other healthcare providers, such as auxiliary nurse/midwives, community midwives, village midwives and other health visitors, may also have acquired appropriate skills *if they have been specially trained*. These individuals frequently form the backbone of maternity services at the periphery, and pregnancy and labour outcomes can be improved by making use of their services, especially if they are supervised by well trained midwives.

Box 3. Cont'd - Who is a Skilled Attendant?

In developed countries and in many urban areas in developing countries, skilled care at delivery is usually provided in a health facility. However, birth can take place in a range of appropriate places, from home to tertiary referral center, depending upon availability and need, and WHO does not recommend any particular setting. Home delivery may be appropriate for a normal delivery, provided that the person attending the delivery is suitably trained and equipped¹ and that the referral to a higher level of care is an option.

¹ In many countries, TBAs have received training in order to promote safer birth practices, including clean delivery and avoidance of harmful practices. However, to fulfill all the requirements for management of normal pregnancies and births and for identification and management or referral of complications, the education, training, and skills of TBAs are insufficient. Their background may also mean that their practices are conditioned by strong cultural and traditional norms, which may also impede the effectiveness of their training.

Source: WHO/UNFPA/UNICEF/The World Bank, Joint statement on Reduction of Maternal Mortality, 1999.

3. Emergency Obstetric Care

Emergency obstetric care (EmOC) refers to a series of crucial life-saving functions, ideally performed in a medical facility, which can prevent the death of a woman experiencing the start of complications during pregnancy, delivery, or the post-partum period. EmOC is a medical response to a life-threatening condition and is not a standard for all deliveries. EmOC functions are often divided into two categories: (1) **basic EmOC**, which can take place at a health centre and be performed by a nurse, midwife or doctor, and (2) **comprehensive EmOC**, which usually requires the facilities of a district hospital with an operating theatre. The essential functions are listed in **table 1** below:

Table 1. Basic and Comprehensive EmOC Functions

Basic EmOC Functions Performed in a health centre without operating room	Comprehensive EmOC Functions Requires an operating room and is usually performed in district hospitals
<ul style="list-style-type: none"> ■ Intravenous antibiotics ■ Intravenous oxytocics ■ Intravenous anticonvulsants ■ Manual removal of placenta ■ Assisted vaginal delivery ■ Removal (by aspiration) of retained products 	All six Basic EmOC functions plus: <ul style="list-style-type: none"> ■ Caesarean section ■ Blood transfusion

The basic EmOC functions consist of administering medications by injection. These are usually antibiotics to treat an infection, anticonvulsants to treat a seizure, or oxytocics to treat excessive bleeding by helping the uterus to contract. Assisted vaginal delivery refers to the use of a vacuum extractor preferably to the use of forceps. A placenta that has failed to be expelled naturally can cause both excessive bleeding and infection. The same is true for retained products of incomplete miscarriage or abortion. Removal of placenta can usually be done manually. Removal of retained products can be done under light anaesthesia and usually requires a minor surgical procedure like a manual vacuum aspiration.

Comprehensive EmOC refers to the ability to perform more complex surgical interventions such as a caesarean section to relieve obstructed labour. It also refers to the ability to administer a blood transfusion to treat life-threatening haemorrhage. Blood must be safely collected, screened and stored; therefore, a complete blood bank is required.

Improving the availability of services is a crucial first step to increase access to EmOC. In many cases only limited inputs are needed to expand existing health facilities and enable them to provide EmOC services. These interventions may include: renovating existing operating theatres or equipping new ones; repairing or purchasing surgical and sterilization equipment; training doctors and nurses in life-saving skills; and improving health services management. Health service management improvements include adequate staffing of health facilities, a steady supply of drugs and other supplies, maintenance of the health infrastructure and equipment, a system allowing 24-hour readiness, and fair health-care service pricing policies. It also means promoting monitoring and evaluation, and constant improvement in the quality of services.

IV. Using Output and Outcome Indicators to Monitor Progress

Introduction

It is difficult to determine whether maternal mortality programme interventions have been successful, as impact indicators such as maternal mortality rates and ratios are often unavailable. Reasons for this unavailability include the poor quality of vital statistics reported by many developing countries and the fact that, when recorded, maternal deaths are often not distinguished from deaths by

other causes⁵. It is therefore recommended that programmes rely on internationally agreed upon indicators: the MDG indicator of skilled attendance at birth and the six "UN EmOC process indicators" agreed upon by UNICEF, WHO and UNFPA⁶. These indicators describe the functionality of health services and the capacity of health systems to address life-threatening complications arising during pregnancy and delivery. It is recommended to also use behaviour change and policy-related indicators to monitor the demand for EmOC and the policy environment.

Current experiences in using the UN Process Indicators in Malawi⁷ concluded that although the UN EmOC Process Indicators have limitations, this monitoring system has provided information vital to health providers, managers and policy makers that enabled them to increase the availability, distribution and quality of services.

Box 4 Highlights experiences in using EmOC process indicators for obstetric service baseline assessments.

Box 4. Findings from obstetric service baseline assessments

In 2000-2001, UNFPA country offices in Cameroon, India, Morocco, Mozambique, Nicaragua, Niger, and Senegal carried out assessments of obstetric services using EmOC process indicators.

Common trends emerge from these surveys: for instance, the real challenge is to expand availability of basic EmOC facilities; the geographic distribution of facilities is skewed and much more effort is needed to make services accessible in rural areas, a problem often compounded by poor roads and lack of transportation; and the case fatality rate cannot be used alone as an indicator of poor quality of services at the facility. Late arrival to the facility rather than quality of services could be the reason for a maternal death.

Source: International Journal of Obstetrics and Gynecology (IJGO) in 2002 and 2003, UNFPA /AMDD Making Safe Motherhood a Reality in West Africa - Using Indicators to Programme for Results. 2003.

The DOPA⁸ indicators and corresponding means of verification (MOVs) outlined in **figure 1** and **table 2** have become important tools to monitor UNFPA's contribution to reducing maternal mortality. For each indicator, a precise definition of how the indicator is constructed, the minimum and/or maximum levels required, and the sources of data used are provided on the following pages.

⁵ Other reasons include: (1) estimates of maternal mortality are based on measurement of these ratios in samples of the population through expensive surveys with wide confidence intervals; (2) only retrospective data can be obtained so it is difficult to measure recent progress; and, (3) maternal mortality ratios may provide an overall national picture but lacks sufficient detail for local level decision-making.

⁶ Maine, Deborah et al. Guidelines for Monitoring the Availability and Use of Obstetric Services. UNICEF, WHO, UNFPA. August 1997.

⁷ Hussein J, Goodburn E A, Damisoni H, Lema V and Graham W. (2001). Monitoring obstetric services: putting the 'UN Guidelines' into practice in Malawi: 3 years on." *Int'l Journal of OB & Gyn* 75, 63-73.

⁸ **DOPA**: Direct, Objective, Practical & Adequate. For further explanation, see The Programme Manager's Monitoring and Evaluation Toolkit Tool 6, Part I: Programme Indicators-The Basic Concepts.

Table 2. Indicators for monitoring EmOC in Maternal Mortality Reduction programmes

Indicator	Optimal Levels
1. Proportion of deliveries assisted by skilled health personnel ^a	According to ICPD: 60% of deliveries According to the MDGs: 90% of deliveries There is usually a national target.
2. Amount of Basic and Comprehensive EmOC facilities available per population	For every 500,000 population , there should be: - At least 4 Basic EmOC facilities - At least 1 Comprehensive EmOC facility
3. Geographical distribution of EmOC facilities (sub-indicators: time to reach EmOC facility and proportion of households within 2 hours of Basic EmOC facility)	Ideally, basic EmOC facilities should be located so they can be accessed within a maximum of 2 hours. Comprehensive EmOC facilities should be accessible within a maximum of 12 hours.
4. Proportion of all births in Basic and Comprehensive EmOC facilities	At least 15% of all births in the population should take place in basic or comprehensive EmOC facilities
5. Met need for EmOC: Proportion of women with obstetric complications who are treated in EmOC facilities	100% of women with obstetric complications should be treated in EmOC facilities.
6. Caesarean sections as a proportion (%) of all births	Caesarean sections should account for no less than 5% and no more than 15% of all births (C-sections performed for emergency purposes only)
7. Obstetric Case Fatality Rate	The case fatality rate among women with obstetric complications in EmOC facilities should be less than 1% (indicator best interpreted at facility level)

Source: Except for the first indicator; the others are adapted from Maine, Deborah et al. *Guidelines for Monitoring the Availability and Use of Obstetric Services*. UNICEF, WHO, UNFPA. August 1997

Table 3 provides an overview of data required to construct the indicators.

Table 3. Types of Data Used to Construct Indicators

Type of Data	Indicator 1	Indicator 2	Indicator 3	Indicator 4	Indicator 5	Indicator 6	Indicator 7
Population Size	•	•	•	•	•	•	
Birth Rate	•			•	•	•	
Number of births assisted by skilled birth attendants	•						
Health Facility Data: EmOC signal functions		•	•	•			
Number of births	•			•		•	
Number of complicated cases in EmOC facilities					•		•
Number of C-sections						•	
Number of maternal deaths (direct causes)							•

Source: Adapted from *Distance Learning Courses on Population Issues: Course 6, Module 2*. UNFPA, 2002.

^a This indicator, proposed to monitor the MDG No.5, is not part of the 6 UN EmOC process indicators originally proposed in the referenced source.

Indicator 1: Proportion of deliveries assisted by skilled birth attendants

Indicator 1, which is not included in the six "UN EmOC process indicators", should be used to report on the Millennium Development Goal of reducing maternal mortality at both global and national levels. It is irrelevant whether the delivery has taken place at home or in a health facility. It may be difficult to collect accurate data regarding skilled attendance from the community due to recall bias (women responding to surveys may have difficulty identifying the skills of their attendant and may not know the exact training their attendant had received).

Indicator 1	Definition	Numerator	Denominator	Optimal Level
Proportion of deliveries assisted by a skilled attendant (regardless of the place of delivery)	Proportion of all deliveries assisted by either a qualified midwife, nurse midwife or trained doctor capable of performing the six basic EmOC functions	Number of deliveries assisted by a skilled birth attendant	Total number of expected deliveries in the catchment area in one year (provided by the simple calculation of crude birth rate multiplied by estimated population in the area (based on last census and updates))	According to the MDGs: 90% of deliveries
MOV: Numerator: demographic and health surveys; Denominator: census information				

Indicator 2: Amount of functional Emergency Obstetric Care facilities

It is essential to assess the availability of facilities for a given population in order to determine if they are sufficient. In general, research has shown that 15% of pregnancies will result in life threatening complications. Based on this figure and knowing the number of expected births in a given population, it is easy to determine the number of women expected to need EmOC services. The standard of four basic and one comprehensive EmOC facility per 500,000 persons has been established by observation in several developing countries. Application of this standard may vary according to the population density, the nature of the geographical terrain, the time to reach facilities from scattered homes, and other variables. More important is the qualification of a facility as basic or comprehensive EmOC facility. Clearly a facility can only be considered a "basic EmOC" facility if all six basic functions have been performed in the past three months. Similarly, a facility can only be considered a "comprehensive EmOC" facility if all six, plus the extra two, functions have been performed in the past three months. Use of this service indicator requires periodic investigations to ensure that facilities labelled as basic and comprehensive are actually performing the appropriate functions.

Indicator 2	Definition	Numerator	Denominator	Optimal Level
Amount of Basic EmOC facilities	Number of health facilities having provided the 6 basic EmOC functions in the last 3 months, per 500,000 population	Number of facilities having provided the 6 basic EmOC functions in the last 3 months in a given area	Population of catchment area	4 units per 500,000 population
Amount of Comprehensive EmOC facilities	Number of health facilities having provided the 6+2 EmOC functions in the past three months, per 500,000 population	Number of facilities having provided the 6+2 EmOC functions in the past three months	Population of catchment area	1 unit per 500,000 population
MOV: Numerator: supervision reports; facility surveys; Denominator: census information				

Indicator 3: Geographic Distribution of EmOC facilities

Simply having enough EmOC facilities is not sufficient; their geographic distribution must also be considered. If all comprehensive EmOC facilities are clustered in urban areas, a large number of women—especially those living in rural areas—will be unable to access services in a timely manner. Unlike the other indicators in this document, Indicator 3 can only be measured by performing spatial analysis with the use of a map or an interactive Geographic Information System (GIS).

In many developing countries, the terrain is rough and communications, roads and transportation are poor. Traditionally, distance has been the indicator used to assess physical service accessibility. In actuality, the time it takes to reach an EmOC facility is a more accurate indicator of physical access. Travelling even relatively short distances may take a very long time. Often the journey to a health-care facility is made on foot, horseback or by donkey cart. Therefore, a useful proxy indicator may be the proportion of households within a given travel time for a woman to reach a basic or comprehensive EmOC facility. Optimally, all women should live within two hours of a basic EmOC facility. This number was selected as a maximum limit because haemorrhage, the most rapidly fatal complication of pregnancy, can kill a mother in two hours. In order to save the maximum number of lives, facilities must be able to treat pregnant women within this timeframe. This complication can be treated at a basic EmOC facility, though some cases may need to be referred to a comprehensive facility for blood transfusions. Therefore, an ideal geographic distribution of facilities would ensure that all women live within two hours of a basic EmOC facility and twelve hours of a comprehensive one. This is clearly an ambitious goal, involving improvements in communication and transportation systems and roads.

Indicator 3	Definition	Mode of Measurement	Optimal Level
Geographic distribution of EmOC facilities	Assessment (by map or GIS), or actual measurement, of physical accessibility to EmOC facilities	Spatial analysis conducted with use of GIS, or proportion of households within 2 hours of a basic EmOC facility	Ideally, all basic EmOC facilities are within two hours travel time and comprehensive EmOC facilities are within 12 hours travel time for women of reproductive age
MOV: supervision reports; accreditation meetings; GIS maps			

Indicator 4: Proportion of all births in functional EmOC facilities

This service indicator measures actual utilization of EmOC facilities. Once it is confirmed that appropriate facilities exist, provide the appropriate services (six or eight functions) and are evenly distributed, it must be determined whether patients are, in fact, utilizing those services. If 15% of women are estimated to experience complications, then at least 15% of births should be taking place in EmOC facilities. Obviously, this crude indicator does not allow for the assessment of which births take place in EmOC facilities. It is conceivable that only non-complicated births are taking place in EmOC facilities and that all complicated ones take place in homes or elsewhere. This indicator should therefore be combined with the indicator of met need for EmOC explained below.

Indicator 4	Definition	Numerator	Denominator	Optimal Level
Proportion of deliveries taking place in EmOC facilities	Proportion of all deliveries taking place in functional EmOC facilities	Number of births taking place in functional EmOC facilities in the catchment area within one year	Total number of expected deliveries in the catchment area in one year	At least 15% take place in a EmOC facility (hoping to "catch" the maximum proportion of complicated cases...)
MOV: Numerator: demographic and health surveys; health service survey; health MIS; Denominator: census information				

Indicator 5: Met Need for EmOC

Met need for EmOC means ensuring that all women with complications are appropriately treated. The goal is that all (100%) women who experience complications are treated at the appropriate level of care. Simply establishing that at least 15% of births are taking place in EmOC facilities does not ensure that all women with complications are being served. Mechanisms should be in place at all EmOC facilities to record (a) whether a woman was actually experiencing a complication, and (b) the type and severity of that complication. The UN Guidelines of 1997 offer a list of seven complications that must be adhered to when assessing this indicator. Some women may

choose to have normal deliveries in EmOC facilities, so the percentage of all births taking place in EmOC facilities may include both normal and complicated deliveries. The percentage of complicated deliveries among those births will vary between rural and urban populations, and at public and private facilities.

Indicator 5	Definition	Numerator	Denominator	Optimal Level
Met need for EmOC	Proportion of women with complications who are treated in EmOC facilities	Number of women admitted to EmOC facilities with one or more of the seven complications described in the UN Guidelines of 1997	Total number of expected deliveries with complications (calculated as 15% of expected births in the catchment population)	All (100%) women with obstetric complications are treated in EmOC facilities
MOV: Numerator: health MIS; maternity admission registers; Denominator: census information				

Indicator 6: Proportion of Caesarean Sections

The proportion of Caesarean sections is a useful service indicator for many reasons. One is that it is likely that C-sections will be adequately recorded in hospital records. Studies indicate that 5% of all births will have complications (e.g. obstructed labour) that require a C-section to ensure maternal survival. A minimum of 5% of births should, therefore, be performed by C-section. This is not an infallible measure, however. In many countries, C-sections are performed in the absence of maternal life-threatening complications for reasons related to the newborn, or for profit, patient preference or hospital protocol. It is important to examine hospital records to determine the number of C-sections performed on women who were experiencing complications. To ensure that C-sections are not performed needlessly (since non-necessary operations carry a risk and have consequences for future births), a maximum level of 15% of all deliveries has been established as a standard.

Indicator 6	Definition	Numerator	Denominator	Minimum/ Maximum Level
Proportion of C-sections	Proportion of C-sections to all births in the population	Number of C-sections in all EmOC facilities in the catchment population in one year	Total number of expected deliveries in the catchment area in one year	At least 5% and not more than 15% of all deliveries
MOV: Numerator: demographic and health surveys; health MIS (facility records); health service surveys; Denominator: census information				

Indicator 7: Obstetric Case Fatality Rate

The final standard service indicator is the obstetric case fatality rate at EmOC facilities. This is a measure of the quality of services at each facility. It is not calculated only for comprehensive EmOC facilities.

It is measured as the number of women with pregnancy-related complications who die in an EmOC facility divided by the number of women with an obstetric complication treated at that facility. In order to obtain a national or regional obstetric case fatality rate, it is necessary to aggregate the data provided by each EmOC facility. In large hospitals, it is possible to disaggregate the obstetric case fatality rate for each type of complication (each complication carries a different type of treatment, which can be assessed separately). Ideally, each facility should have an obstetric case fatality rate of 1% or less.

This measure is most useful to track progress in the quality of services within a certain facility over time. However, it does not take into account the condition of the patients upon arrival at the facility. This makes it difficult to make comparisons among facilities in drastically different locations, or those that serve dramatically different populations. Careful interpretation of facility records is necessary if record keeping at the comprehensive EmOC facility is poor. Additionally, one should be aware that the obstetric case fatality rate may be low if it is practice at the given facility to send women with complications home to die, or if women with severe complications are transferred to intensive care units and lost to follow-up.

Indicator 7	Definition	Numerator	Denominator	Optimal Level
Obstetric case fatality rate in EmOC facilities	Proportion of women with an obstetric complication who die in EmOC facilities	Number of direct obstetric deaths in EmOC facility(ies) in one year	Number of obstetric complications in the same facility(ies) in one year	Obstetric case fatality rate should be less than 1%
MOV: Numerator: facility service statistics; maternal mortality audit; Denominator: census information				

A number of public health researchers have questioned the relevance of the indicator of "skilled attendance at birth" and the UN EmOC process indicators to track maternal mortality. **Box 5** provides a summary of some of their concerns.

Box 5. Process Indicators for EmOC: How Useful Are They?***Addressing Utilization***

Proportion of births attended by skilled health personnel: While this indicator reflects national trends in access to skilled care at birth, it does not indicate which specific components of the health system need strengthening. Is it the care provided on the spot, at home or at the first referral level, or at the second referral level? It is also difficult to obtain information on the "skills" of the birth attendant when interviewing patients or relatives during community-based surveys.

Proportion of C-sections: Population-based estimates of the proportion of C-sections performed may reflect the extent to which pregnant women access EmOC services. However, as C-section proportions rise, it may be possible that the majority of these deliveries are performed to avoid problems, whether they truly exist or not. It would be important to differentiate C-sections performed in emergency from those performed for convenience.

Proportion of births in EmOC facilities: The 1997 joint UNICEF/WHO/UNFPA guidelines suggest that at least 15% of all women should deliver in basic and comprehensive EmOC facilities. While this indicator can be useful in determining utilization, the numerator may contain women with a normal delivery, and not necessarily those experiencing emergency obstetric complications. Further, the assumption that 15% of pregnant women are bound to experience obstetric emergencies is not supported by empirical evidence.

Addressing Met Need

Proportion of all women with complications who are treated in EmOC facilities: This indicator has widely been accepted as an indicator of "met need." However, before using this indicator, the following four issues must be addressed: (1) it is necessary to define "complications"; (2) while abortion and ectopic pregnancy may be important causes of maternal death, they are more difficult to incorporate in the list of obstetric complications because they tend to appear in the earlier stages of pregnancy; (3) it has never been empirically verified that 15% of all births are "complicated," nor is there any reason to believe that the incidence of obstetric complications is constant across population groups (see above); and, (4) a limitation of this indicator is the assumption that EmOC for the broad range of complications specified can only be delivered in health facilities. If skilled attendants are present during home births, they may perform basic functions, which will prevent death, and contribute to a decline in maternal mortality rates.

Source: adapted from Ronsmans C, Campbell O, Mc Dermott J and Koblinsky M (2002) "Questioning the indicators of need for obstetric care" Bulletin of the World Health Organization, 80(4) 317-324.

Sources

Hussein J., Goodburn E. A., Damisoni H., Lema V. and Graham W. (2001). **"Monitoring obstetric services: putting the 'UN Guidelines' into practice in Malawi: 3 years on"**, Int'l Journal of Ob & Gyn 75, 63-73.

International Confederation of Midwives Website: <http://www.internationalmidwives.org/>

Maine, Deborah; Murat, Akalin Z.; Ward, Victoria M.; Kamara, Angela. **"The Design and Evaluation of Maternal Mortality Programs. Centre for Population and Family Health"**, Joseph L. Mailman School of Public Health, Columbia University. June 1997. This is a very comprehensive guide for planning and monitoring maternal mortality reduction programmes and is accessible on the Internet at www.amdd.hs.columbia.edu

Maine, Deborah et al. **"Guidelines for Monitoring the Availability and Use of Obstetric Services"**, UNICEF, WHO, UNFPA. August 1997. www.unicef.org/reseval/pdfs/finalgui.pdf

Population Reference Bureau. Yvette Collymore, **"Improving Emergency Care to End Maternal Deaths"**, PRB, November 2003.
http://www.prb.org/Template.cfm?Section=PRB&template=/Content/ContentGroups/Articles/03/Improving_Emergency_Care_to_End_Maternal_Deaths.htm

Safe Motherhood Website: <http://www.safemotherhood.org/index.html>

UNFPA Website: www.unfpa.org

UNFPA. **"Maternal Mortality Update 2002: A Focus on Emergency Obstetric Care"**, UNFPA, 2003. http://www.unfpa.org/upload/lib_pub_file/201_filename_mmupdate-2002.pdf

UNFPA. **"Using Indicators to Programme for Results: Making Safe Motherhood a Reality in West Africa"**, UNFPA, 2003.
http://www.unfpa.org/upload/lib_pub_file/149_filename_safemwestfrica.pdf

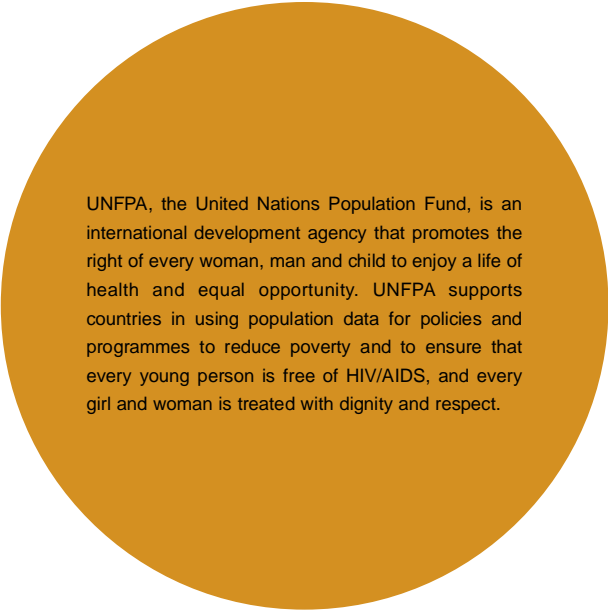
UNFPA/AMDD. **Distance Learning Courses on Population Issues, Course No.6, Reducing Maternal Death: selecting priorities, tracking progress**, 2002.

Designed by: 

Published:

Programme Manager's Planning Monitoring & Evaluation Toolkit

Size: 21x29.7cm * Copies 200 * License No. 59-2008/CXB/260-79/GTVT, dated 04/12/2008



UNFPA, the United Nations Population Fund, is an international development agency that promotes the right of every woman, man and child to enjoy a life of health and equal opportunity. UNFPA supports countries in using population data for policies and programmes to reduce poverty and to ensure that every young person is free of HIV/AIDS, and every girl and woman is treated with dignity and respect.