HANDBOOK OF INDICATORS FOR FAMILY PLANNING PROGRAM EVALUATION

Bertrand, Magnani, Knowles

HANDBOOK OF INDICATORS FOR FAMILY PLANNING PROGRAM EVALUATION

■

Jane T. Bertrand
Robert J. Magnani
Naomi Rutenberg

The EVALUATION Project

December 1994

USAID Contract Number: DPE–3060–C–00–1054–00

Acknowledgments

We wish to acknowledge the support and contribution from the Office of Population, United States Agency for International Development (USAID Contract Number: DPE–3060–00–C–1054–00) and thank them for making this publication possible.

The authors wish to express their thanks to the many members of the international population community who have participated in the working groups organized under The EVALUATION Project. These meetings have served to critically analyze and subsequently to build consensus on many of the indicators listed in this Handbook. The names of the individuals participating in these groups are listed in the minutes of the different meetings. Though this group is too numerous to mention by name in this section, we sincerely appreciate the time and effort you have invested in assisting in this process.

We are also grateful to several individuals who have contributed to the preparation of drafts used in the handbook: Barbara Seligman and Marc Mitchell (section on Management): Lynne Gaffikin and Lisanne F. Brown (Training); John Stover, Susan Smith, and Rich Owens (Commodities and Logistics); Phyllis T. Piotrow, D. Lawrence Kincaid and Young–Mi Kim (Information–Education–Communication); and Naomi Rutenberg (Demand for Family Planning, Contraceptive Practice, and Fertility Impact).

We thank Ms. Myrna Seidman for her ideas regarding the layout used in presenting the indicators, Mr. Trevor Croft for providing figures for some of the illustrative computations, and Ms. Kimberly Meyer–Ramirez for her assistance in compiling the text. For the design and layout of the final document, we extend our thanks to Sue Koenigshofer (SCK Design), Laura Kline (the Carolina Population Center), and John Jaszczak (Data Spectrum Inc.).

Several individuals have reviewed and provided valuable feedback on one or more drafts of the Handbook. External reviewers include: Jay Bainbridge, R.A. Bulatao, Snehendu B. Kar, Mark Laskin, W. Parker Mauldin, John Ross, Martin Vaessen, and Ann Way. Members of The EVALUATION Project staff include: Krista Stewart (Technical Advisor, USAID / Washington); Amy Tsui, Al Hermalin, Sharon Kirmeyer, Linda Lacey, Naomi Rutenberg, Krishnamurthy Srinivasan, John Stover, and James Veney.

In extending thanks to all who have participated in this process, the authors take full responsibility for the content of this document.

TABLE OF CONTENTS

	Summary of List of Indicators in the Handbook	4
	List of Acronyms	8
Chapter		
I	Overview of the Handbook	10
II	Indicators to Measure the Policy Environment	24
III	Indicators to Measure Service Delivery Operation	46
	A. Management B. Training C. Commodities and Logistics D. Information – Education – Communication (I–E–C) E. Research and Evaluation	
IV	Indicators to Measure Family Planning Service Outputs	100
	A. Accessibility (Illustrative) B. Quality of Care (Illustrative) C. Program Image	
V	Indicators to Measure Demand for Children (Fertility Demand)	122
VI	Indicators to Measure Demand for Family Planning	130
VII	Indicators to Measure Service Utilization	142
VIII	Indicators to Measure Contraceptive Practice	154
IX	Indicators to Measure Fertility Impact	168
	A. Fertility Level B. Births Averted C. Other Indicators	
X	Future Steps in the Development of Indicators for Family Planning Evaluation	188
	References	192
Appendix		
A	Listing of All Indicators by Primary Source of Data	200
B	Indicators for Evaluating the Management Functional Area	206
C	Indicators of Quality of Care in Family Planning Programs	209
D	Indicators for Evaluating Program Use of Operations Research	215
E	Illustration of the Calculation of Unmet Need	217

SUMMARY LIST OF INDICATORS

Chapter	Indicators

II Indicators to Measure the Policy Environment

- Existence of a policy development plan
- Number of appropriately disseminated policy analyses
- Number of awareness–raising events targeted to leaders
- Existence of a strategic plan for expanding the national family planning program
- Integration of demographic data into development planning
- Number of statements of leaders in support of family planning
- Formal population policy addressing fertility and family planning
- National family planning coordination
- Level of the family planning program within the government administration
- Levels of import duties and other taxes
- Restrictions on advertising of contraceptives in the mass media
- Absence of unwarranted restrictions on providers
- Absence of unwarranted restrictions on users
- Public sector resources devoted to family planning as a percentage of GDP
- Quality of program leadership
- Extent of commercial sector participation

III Indicators to Measure Service Delivery Operations

A. Management

- Existence of a clear mission that contributes to the achievement of program goals
- Realization of operational targets
- Clearly defined organizational structure
- Adequacy of staffing
- Awareness of current financial position
- Access to current information in key areas of program functioning
- Access to current information on program progress
- Capacity to track commodities

B. Training

- Number/percentage of courses that achieve learning objectives
- Number/percentage of courses that contribute to the achievement of program training objectives
- Number/percentage of courses where the training methodology is appropriate for the transfer of skills and knowledge
- Number of trainees by type
- Number/percentage of trainees who have mastered relevant knowledge

- Number/percentage of trainees competent to provide a specific family planning service
- Number/percentage of trained providers assessed to be competent at a specific period (e.g., 6 months) post–training
- Number/percentage of trainees who apply the skills to their subsequent work

C. Commodities and Logistics

- Pipeline wastage
- Percentage of storage capacity meeting acceptable standards
- Frequency of stock–outs
- Percentage of service delivery points (SDPs) stocked according to plan
- Percentage of key personnel trained in contraceptive logistics
- Composite indicator for commodities and logistics

D. Information–Education–Communication (I–E–C)

- Number of communications produced, by type, during a reference period
- Number of communications disseminated, by type, during a reference period
- Percentage of target audience exposed to program messages, based on respondent recall
- Percentage of target audience who correctly comprehend a given message
- Number of contraceptive methods known
- Percent of audience who acquire the skill to complete a certain task as a result of exposure to a specific communication
- Percentage of target audience exposed to a specific message who report liking it
- Number/percentage of target audience who discuss message(s) with others, by type of person
- Percentage of target audience who advocate family planning practice

E. Research and Evaluation

- Presence of an active research and evaluation unit
- Extent of use of a service statistics system
- Conduct of periodic household and/or special purpose surveys and studies
- Conduct of operations research (OR)
- Regular conduct of process evaluations
- Conduct of effectiveness, efficiency, and impact evaluations
- Use of research and evaluation results for program modification
- Dissemination of research and evaluation results

IV Indicators to Measure Family Planning Service Outputs

A. Accessibility (Illustrative)

- Number of SDPs located within a fixed distance or travel time of a given location (i.e., service density)
- Cost of one month's supply of contraceptives as a percentage of monthly wages

- Restrictive program policies on contraceptive choice
- Percentage of the population who know of at least one source of contraceptive services and/or supplies
- Percentage of non–use related to psycho–social barriers

B. Quality of Care (Illustrative)

- Number of contraceptive methods available at a specific SDP
- Percentage of counseling sessions with new acceptors in which provider discusses all methods
- Percentage of client visits during which provider demonstrates skill at clinical procedures, including asepsis
- Percentage of clients reporting "sufficient time" with provider
- Percentage of clients informed of timing and sources for resupply/revisit
- Percentage of clients who perceive that hours/days are convenient

C. Program Image

- Number and type of activities to improve the public image of family planning during a reference period (e.g., one year)
- Percentage of target population favorable to the (national) family planning program

V Indicators to Measure Demand for Children (Fertility Demand)

- Mean desired family size
- Desire for additional children
- Desired status of previous births
- Wanted total fertility rate (WTFR)

VI Indicators to Measure Demand for Family Planning

- Demand for limiting
- Demand for spacing
- Total demand (for family planning)
- Unmet need for family planning
- Satisfaction of demand for family planning

VII Indicators to Measure Service Utilization

- Number of visits to service delivery point(s)
- Number of acceptors new to modern contraception
- Number of acceptors new to the institution
- Number of new segment acceptors

- Couple years of protection (CYP)
- Method mix
- User characteristics
- Continuation rates

VIII Indicators to Measure Contraceptive Practice

- Contraceptive prevalence rate (CPR)
- Number of current users
- Level of ever (past) use
- Source of supply (by method)
- Method mix
- User characteristics
- Continuation rates
- Use failure rates

IX Indicators to Measure Fertility Impact

A. Fertility Level

- Crude birth rate (CBR)
- Age–specific fertility rate (ASFR)
- Total fertility rate (TFR)

B. Births Averted

- Births averted (by the program)

C. Other Indicators

- Parity–specific birth rate
- Proportion of births above (or below) a specified parity
- Proportion of births by women above or below a specified age
- Median length of birth intervals
- Proportion of open or closed birth intervals that are of a specified length or longer
- Unwanted total fertility rate (UTFR)

LIST OF ACRONYMS

Technical Acronyms

ASFR	Age–Specific Fertility Rate
CBD	Community–Based Distribution
CBR	Crude Birth Rate
CPR	Contraceptive Prevalence Rate
CPS	Contraceptive Prevalence Survey
CYP	Couple–years of Protection
DHS	Demographic and Health Surveys
DTFR	Desired Total Fertility Rate
FEFO	First–Expiry, First–Out
FIFO	First–In, First–Out
FP	Family Planning
FPS	Family Planning Survey
GDP	Gross Domestic Product
I–E–C	Information–Education–Communication
IUD	Intrauterine Device
KAP	Knowledge, Attitude, Practice
LAM	Lactational Amenorrhea Method
LMIS	Logistics Management Information System
L&M	Lapham and Mauldin Family Planning Program Effort Measures
MASFR	Marital Age–Specific Fertility Rate
MCH	Maternal and Child Health
MIS	Management Information Systems
NFP	Natural Family Planning
OR	Operations Research
PSBR	Parity–Specific Birth Rate
SDP	Service Delivery Point
SCYP	Standard Couple–Years of Protection
TA	Technical Assistance
TFR	Total Fertility Rate
UTFR	Unwanted Total Fertility Rate
VSC	Voluntary Surgical Contraception
WFS	World Fertility Survey
WTFR	Wanted Total Fertility Rate

Organizational Acronyms

APROFAM	Asociación ProBienestar de la Familia (Guatemala)
AID	Agency for International Development
CA	Cooperating Agencies
CDC	Centers for Disease Control
CEDPA	Center for Development and Population Activities
FHI	Family Health International
IPPF	International Planned Parenthood Federation
NGO	Non–Governmental Organization
RAPID	A.I.D. funded – Resources for the Awareness of Population in Development Project
SDWG	Service Delivery Working Group of The EVALUATION Project
USAID	United States Agency for International Development

Chapter I

Overview

- Introduction
- Objectives of the Handbook
- Intended Audience
- Contributors to this Handbook
- Use of the Handbook
- Organization of the Handbook
- Overview of the Conceptual Framework
- Types of Indicators: Input, Process, Output, and Outcome
- Sources of Data
- Scoring of Indicators
- Anticipated Update of this Handbook

Chapter I

Overview of the Handbook

More evaluation work has been done in the field of family planning than for any other type of public health or social intervention. This effort has resulted from a sustained commitment on the part of governments, international donors, and individual researchers over several decades to understand the mechanisms that explain contraceptive use and ultimately fertility reduction. More recently, interest in the contribution of family planning to improved maternal and child health outcomes and in its importance as a basic human right have provided further impetus to assess the effectiveness of these programs.

Much of the early work in family planning evaluation, dating to the 1960s and 1970s, focused on particular outcomes, specifically contraceptive use and fertility rates. However, as the field has advanced, considerable attention has also been given to the different components that comprise the overall program and to performance in these functional areas. As a result of three decades of extensive work in this field, a rich reserve of indicators now exist to measure family planning program performance and impact. These can be found in the voluminous literature on family planning evaluation. Key concepts and definitions are summarized in a United Nations publication on the methodology of family planning program evaluation (United Nations, 1986).

Despite this wealth of experience and extensive documentation of family planning evaluation, there are two major shortcomings with regard to indicators in this field. First, the definitions of indicators used to evaluate family planning programs lack consistency. Second, the diverse indicators available from the family planning literature have never been compiled into a single, readily accessible and user–friendly source. This Handbook is designed to address these two shortcomings.

Objectives of the Handbook

The Handbook provides a comprehensive listing of the most widely used indicators for evaluating family planning programs in developing countries. Moreover, the indicators are organized according to the conceptual framework developed under The EVALUATION Project, which maps the pathways through which programs achieve results and thus provides a logical framework for developing an evaluation plan.

Whereas some past evaluation efforts have tended to treat the operations of the family planning program as a "black box," this framework specifies how programs are expected to achieve results at both the program level and the population level. Thus, in addition to determining whether the program achieved the desired outcome at the population level, one can also trace through the different components of the system to identify strong points to be reinforced and shortcomings to be redressed.

The specific objectives of this Handbook are:

- to differentiate the levels for which family planning evaluation indicators are needed (program versus population level);
- to compile in a single volume those indicators judged most useful in family planning program evaluation;
- to provide a definition of these indicators in an effort to enhance the consistent use of terms across programs, countries, and donor agencies; and
- to promote evaluation of programs by making indicators better known and easier to use.

Intended Audience

Several different audiences should find this Handbook very pertinent to their own work, including:

1. Administrators/managers in family planning programs worldwide
   - to assess whether the evaluations done by their own staff or by external groups include the key indicators for assessing program performance.
2. Staff in international family planning agencies responsible for designing and evaluating collaborative projects with host country institutions
   - to compare definitions of indicators currently in use with the operational definitions included in this Handbook to ensure consistency in terms;
   - to choose from this "menu" of possible indicators in developing an evaluation plan for ongoing and/or future activities;
   - to identify the output that can be expected from different functional areas and that can be reasonably linked to the activities conducted in that area; and
   - to recognize the difficulty of linking activities in functional areas directly to population–based effects (e.g., contraceptive prevalence), except where an experimental or quasi–experimental design is used.
3. In–country evaluation specialists responsible for monitoring performance and evaluating the effectiveness of family planning programs in specific settings
   - (same as for the staff of international family planning agencies, directly above); and
   - to recognize important caveats in the use of specific indicators, which may be cited in the presentation of evaluation results.
4. Applied family planning researchers and demographers interested in fertility change
   - to use the conceptual framework described herein as a point of reference in designing operations research and other types of applied research projects; and
   - to expand the traditional approach to the analysis of fertility decline, with its strong emphasis on structural determinants of fertility demand, to include a more detailed examination of the family planning supply environment and its impact on contraceptive use/fertility decline.

Contributors to this Handbook

Although the actual compilation of this Handbook was done under The EVALUATION Project, numerous parties have contributed to it:

USAID Cooperating Agencies (CAs)

In early 1992, a questionnaire was sent to all CAs in the Office of Population to ascertain the types of indicators that they (or their host–country collaborators) use in evaluating performance and impact. Those indicators that respondents frequently cited are included in this handbook.

Working Groups of Specialists in Functional Areas

This first edition of the Handbook has benefited from the insights of six working groups organized under the auspices of The EVALUATION Project. These groups have been comprised largely of staff from the USAID Office of Population and its cooperating agencies. The work has focused on six functional areas: policy, management, training, commodities and logistics, operations research, and service delivery. Of the multiple facets of service delivery, the Service Delivery Working Group has concentrated primarily on one area to date: quality of care. Thus, the indicators presented in this Handbook for policy, management, training, commodities/logistics, research, and quality of care have benefited from input of various specialists within the USAID population community.

The Handbook also reflects the conclusions of three earlier working groups. Two of these groups were sub–committees convened under an USAID Task Force for Improving Family Planning Program Performance Indicators (operational from 1986–1991):

- the Subcommittee on Quality of Care (Subcommittee on Quality Indicators in Family Planning Service Delivery, 1990); and
- the Subcommittee on Surveys, MIS, and Special Studies (USAID Task Force, 1987).

The third group whose work has contributed to this Handbook was convened by the Information and Training Division of the Office of Population:

- the Evaluation Working Group (covering Training, Information–Education–Communication,and Management) (Landry, 1992).

Staff of The EVALUATION Project

Senior staff of The EVALUATION Project have contributed to specific sections and/or reviewed the entire document. Their previous experience in family planning evaluation and their use of the indicators in relation to ongoing projects have made their assistance particularly valuable.

Use of the Handbook

The compilation of this inventory of indicators to increase consistency in the definition of terms has generally met with enthusiasm among members of the international population community. Nonetheless, there has been some apprehension over the purpose of such a Handbook. Would CAs (U.S.–based organizations that receive USAID funding to promote and conduct family planning activities in developing countries) feel constrained to alter their mode of doing business to "score well" on specific indicators? Would national programs be compared across countries in ways that would influence budget allocations? Would host country institutions find themselves obliged to use this set of indicators in some standardized fashion that might not meet local needs? Would researchers find their free pursuit of knowledge constrained by a pre–established list of indicators?

This Handbook has been compiled as a reference document for use throughout the international population community. It does not give "how to" instructions for planning and implementing evaluations.[1] Rather, it provides *a menu of indicators to be used selectively as part of the evaluation of national family planning programs, regional programs, or projects.*

It should be stressed that no program should attempt to use *all* the indicators outlined in this Handbook. In fact, for routine monitoring purposes it is desirable to select a few relevant indicators that are easy for staff to collect, interpret, and discuss. Special studies can then be conducted to evaluate how programs are doing in areas of particular interest to program staff (staggered to minimize the research burden on the organization).

The lack of consistency in the definition of indicators and terms that exists within the international population community has arisen NOT from a conscious decision to deviate from standard usage, but rather from lack of a central source for obtaining a "standard definition." This problem occurs more frequently with respect to indicators of process and output in family planning programs, such as "new acceptor," than it does with demographic measures based on Demographic and Health Survey data.

Some family planning organizations may find one or more of the definitions contained in this Handbook at odds with their own operational definition of these terms. This situation would require a conscious decision as to the most appropriate action to take. If the cost of making the change is low, the organization might well opt to conform to a standard definition of terms. In some cases, however, there may be strong reasons to adhere to existing definitions in specific settings. Even if an organization opts not to accept these definitions, it would do the international population community a service by explaining (in reports intended for circulation outside the institution) how its own definition of terms differs from that given in this Handbook.

Similarly, researchers are by no means bound by the definitions used herein; others, including "hybrids," may be equally valid in specific settings. Yet the very existence of a standardized set of indicators/definitions will allow greater clarity in those situations where the researcher or program evaluator chooses to define terms differently and offers the rationale for this decision.

In general, however, it is expected that researchers, evaluators, and program managers will either:

- already be using some of the indicators as defined herein; or

[1] Other references are available on the steps entailed in conducting evaluations of family planning programs, including the recent volume by Garcia–Nuñez (1992). In addition, The EVALUATION Project is currently developing a manual on the design and implementation of evaluation plans for national family planning programs that will serve as a companion volume to this Handbook.

- welcome guidance on indicators not yet in use but of potential benefit to the organization or to a specific evaluation.

To the extent that no other definition of an indicator is already in use, there are benefits to adopting a uniform definition that will increase consistency in usage across projects, programs and countries. This uniformity of terms will result in greater understanding of data reported from different sources and more valid cross–country comparisons.

Many have asked, "What are the 10–15 key indicators that are essential for evaluating family planning programs worldwide?" The answer depends entirely on the purpose of the evaluation. A few widely used indicators of outcome (contraceptive prevalence, total fertility rate) would figure in most national evaluation plans. In addition, most programs will monitor at least one output measure, such as new acceptors or couple–years of protection (CYP). However, beyond this, the indicators most important to an evaluation plan will vary according to the objectives of the program, the priority areas within the program, and the availability of data.

The indicators presented in this manual are by no means exhaustive; to the contrary, they barely scratch the surface in some of the functional areas. They are intended to provide a general framework within which a more specific evaluation could be designed. For example, a program manager wishing to obtain feedback on a comprehensive training course covering multiple topics might feel at first glance that the output indicators provided herein were very sketchy, such as:

- number or percent of trainees that have mastered relevant knowledge, and
- number or percentage of trainees who apply skills to their subsequent work.

However, these indicators serve to outline the areas that should be considered in an actual evaluation. Depending on the content of the course, an individual trainee might be queried on 30–50 knowledge items; he/she might be observed on 10, 20, or more tasks in the work place. Because these items would be specific to a given training program, they are not elaborated on in this manual. However, the general categories of indicators to be addressed should serve a purpose in developing evaluation instruments for specific applications.

This Handbook aims to improve program evaluation for the purpose of strengthening family planning service delivery in the national family planning program (regional program, or specific project) of a given country. It is not intended to yield a report card that pits one country against another for future funding consideration. Although this Handbook is designed to encourage the consistent use of definitions and terms across countries or programs, it is important to recognize that the socio–economic and cultural context in which programs operate differs greatly. For example, it is far easier to achieve a certain level of coverage for a family planning media campaign in a small country with a monolingual population, a well developed communications system, and a favorable government policy toward family planning than in a country where these conditions are not present. Similarly, it is important to take into account the magnitude of the achievement in terms of the size of the population in absolute numbers, which would not be apparent from percentages and rates alone. Most would agree that it would be a greater feat to achieve a well–functioning logistics system in India (or even one state in India) than in countries with a similar socio–economic level, but a fraction of the population.

While cross–national comparisons on certain variables are inevitable and often highly useful to the international population community (Baldwin, 1992), the results obtained for these indicators for family planning evaluation are subject to misuse if they are not interpreted in a country–specific context. In many cases it will be far more productive to compare indicators for a given program over time than to use the indicators for cross–national comparisons.

Organization of the Handbook

This Handbook is organized around a conceptual framework that outlines the pathways by which family planning programs achieve impact in a given country.

The term family planning program refers to an organized program often governmental in sponsorship, support, administration, facilities, and personnel, but frequently involving private efforts (family planning associations, private physi-

cians) and commercial ones—designed to provide the information, supplies, and service of modern means of fertility control to those interested (Ross, 1992). In this sense, we include other non–governmental organizations (NGOs) providing contraception under the umbrella of the national program.

Family planning programs tend to have one of three goals: to reduce fertility rates, to improve maternal/child health status, or to enhance the reproductive freedom of women by providing them with means to control their fertility. In some cases the stated objective of the program will be to increase contraceptive prevalence, but this objective is generally motivated by one of the three factors mentioned above.

In this edition of the Handbook, we focus on the first of these objectives (reducing fertility) in detailing the processes by which programs change behavior and achieve impact.[2] It should be noted, however, the conceptual framework presented herein could be readily adapted for use in evaluating family planning programs designed to improve health status or increase reproductive choice.[3]

If one's objective is to reduce fertility, there are a number of factors to take into account (many of which are included as indicators in this Handbook). For example, is there a felt need to space or limit births in this society? Does the population know the available methods for preventing pregnancy? Do they have access to family planning services? Does the family planning program have a positive image that draws new clients? Are the services of sufficient quality to ensure correct and sustained use of contraceptive methods?

This conceptual framework, developed under The EVALUATION Project (Tsui et al., 1992), provides the basis for the organization of the Handbook. The indicators listed herein match the components of the conceptual framework. Specifically, the blocks shown in Figures I–1 and I–2 correspond to the headings of Chapters II–IX, "Indicators to Measure"

- Policy Environment
- Service Delivery Operations
- Service Output
- Demand for Children
- Demand for Family Planning
- Service Utilization
- Contraceptive Practice
- Fertility

Overview of the Conceptual Framework

In any country, there are multiple social, economic, and cultural factors that operate at a societal level to determine the norms of family size. These combine with socio–demographic characteristics and psycho–social factors to influence desired family size at the individual level. Level of socio–economic development, degree of urbanization, the demand for child labor, old–age support and economic security, the cost of raising children, infant mortality rates, levels of female education, status of women, kinship structures, conjugal patterns, and religious customs are all examples of key factors that determine the demand for children in a given society. In the sociology literature, these are referred to as structural determinants; they have worked historically to sustain high levels of fertility in countries at the low end of the development scale.

However, the demand for children is also affected by the family planning supply environment. Many argue that by making family planning services more readily available, one can not only fulfill a latent demand for spacing or limiting that may exist in a given society, but also create a demand for these services by providing couples with alternatives to continued childbearing and the means of achieving pregnancy prevention.

The supply factors in family planning programs, represented as a single block in Figure I–1, are

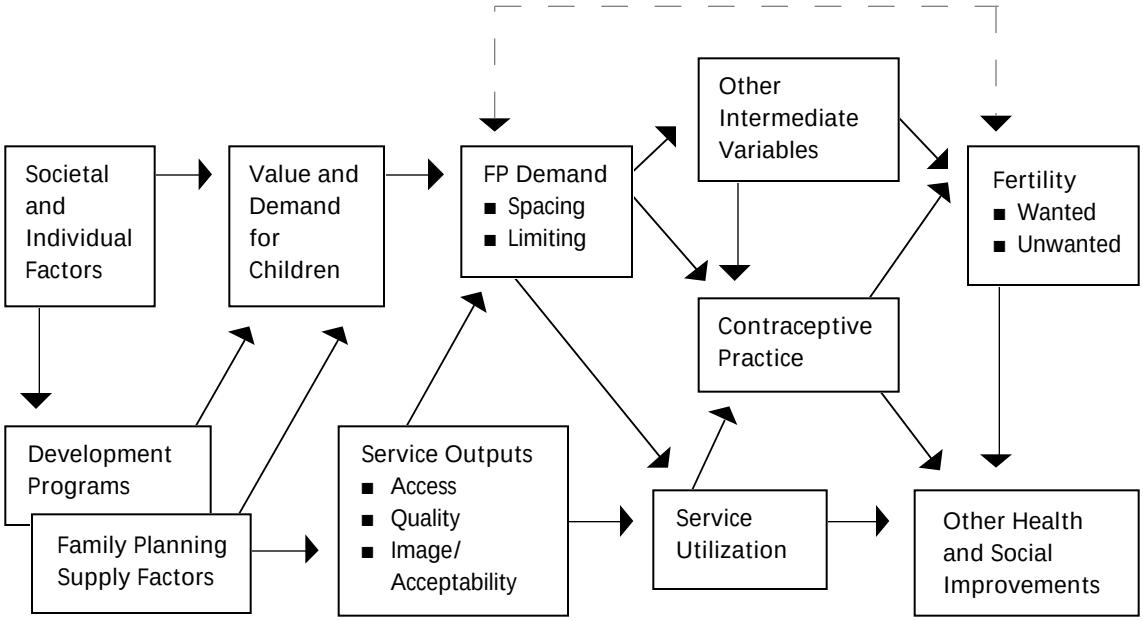
---

[2] The decision to focus on fertility relates to the original objectives of The EVALUATION Project. Whereas these objectives have evolved to include a broader focus on other outcomes, including maternal/child health and reproductive choice, this change is relatively recent. These other outcomes will be covered in the subsequent version of the Handbook, described in Chapter X.

[3] Historically, the primary target for family planning services has been adults (primarily women) in marital union. However, as the focus of programs broadens to include a wider range of reproductive health services, so will the indicators need to be expanded to better reflect the needs of young people not in union, infertile couples, men, commercial sex workers, and others.

Figure I–1

Conceptual Framework of Family Planning Demand and Program Impact on Fertility



shown in more detail in Figure I–2 (see next page). In the vast majority of developing countries, external development assistance has served as a catalyst to the expansion of family planning services (a notable exception is China), affecting both political/administrative systems and the organizational structure.

Key to the development of a strong family planning program are the political and administrative systems within which the program will operate. Political support is widely recognized as a key factor in program success. Resource allocations, the use of provider and acceptor incentives, and cost recovery will determine adequacy of the facilities, staff, equipment, commodities, and other materials needed for a viable program. Legal codes and regulations will affect the flow of contraceptives into the countries, the number of methods that are legal, the characteristics of persons eligible to receive them, and related issues that influence access to contraception. The political and administrative systems for family planning do not exist in a vacuum, but rather are influenced by the larger societal and political governance factors.

A program's policy environment also influences how family planning is organized in a given country: the infrastructure available for service delivery, the extent of integration of family planning with other sectors of the government, the types of service delivery strategies used (clinic–based, community–based distribution, and contraceptive social marketing), and the relative contribution of the public and private sector to the effort.

A comprehensive family planning service delivery program consists of a number of operations (elsewhere called enabling systems, sub–systems, or functional areas). These operations correspond closely to the divisions found in most governmental or private family planning programs: management, supervision, training, commodities and logistics, information–education–communication (I–E–C), and research–evaluation. In short, the input into the family planning program are invested in a set of activities ("processes") that correspond to these functional areas.

What results do programs hope to achieve through this investment of human and financial resources in a defined set of activities? The first result involves the adequacy of the actual service delivery from the client perspective. Specifically, programs strive to improve the quantity and quality of contraceptive services available to the potential client in both the public sector (government program) and private sector (NGO–sponsored services, private doctors, pharmacies or other commercial outlets). Concurrently, through information–education–communication (I–E–C) efforts they attempt to make the social climate for family planning more acceptable, such that clients will receive social support for their decision to contracept.

The adequacy of family planning services is measured in terms of service output. The term output refers to results achieved at the program level; service refers to the objective of improving the service delivery system. Service output can be classified and evaluated on three dimensions:

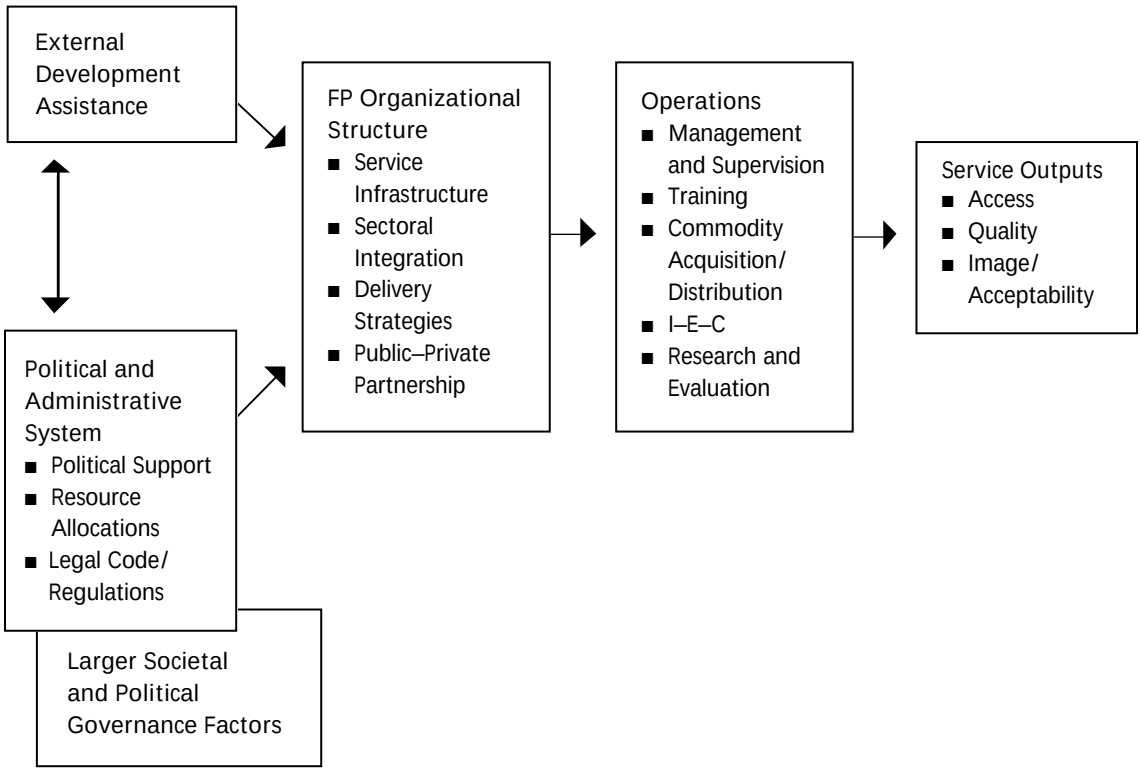
- accessibility of family planning services;
- quality of services; and
- image/acceptability of the program.

By making services more accessible and satisfactory to potential clients, national family planning programs strive to achieve the second key output: an increase in the *utilization of these services*.[4] In many programs, evaluation efforts focus primarily on monitoring indicators of service

[4] In the process, their efforts may affect use in the private sector, if, for example, the I–E–C for the national program has a spin–off effect of increasing commercial pharmacy sales of contraceptive products. These efforts may also affect "non–service" utilization, in the case of individuals who resort to withdrawal or other methods not requiring any type of contact with the program.

Figure I–2

Conceptual Framework of Family Planning Supply Factors



utilization, such as number of new acceptors or couple–years of protection, described in Chapter VII. Indeed, it is often the case that at the program level (i.e., based on data generated from the program itself), the success of the program is evaluated in large part by indicators of service utilization.

In the case of family planning, service utilization is important because it is closely linked with the key behavioral change sought: increased contraceptive practice among the target population. Contraceptive prevalence refers to the percentage of women of reproductive age in the target population (or their partners) using a contraceptive method at a given point in time (often though not always based on women married or in sexual union).

Contraceptive use directly affects fertility. In analyses of the factors responsible for recent fertility decline in developing countries, contraceptive use emerged as the most important. Other intermediate variables (or proximate determinants of fertility) include the percentage of women of reproductive age in sexual unions, the percentage in the non–susceptible post–partum period, and the prevalence of abortion. The other intermediate factors do affect fertility but to a lesser degree (Bongaarts, 1978). (The situation is somewhat different in Africa, where both contraceptive use and post–partum non–susceptibility are key determinants of fertility.)

Contraceptive use also influences MCH outcome, although the exact nature of the relationship is less easy to quantify. Evidence continues to mount on the health benefits associated with (1) avoiding births at too low or too high a maternal age, at high parity, and at short intervals, and (2) reducing reliance on abortion (National Research Council, 1989).

In sum, the titles of the chapters that follow correspond to specific boxes on the conceptual framework. The indicators described in each chapter provide concrete measures for monitoring program performance and measuring outcome in family planning programs.

Types of Indicators: Input, Process, Output, and Outcome

Before proceeding to the definition of actual indicators, we briefly review the terms used in program evaluation. This review is particularly important, given a lack of standardization not only within the family planning field, but also within program evaluation as a science (Veney, 1992). Even those who specialize in evaluation may inadvertently use certain terms interchangeably in informal discussions when, in fact, the terms have a specific technical meaning.

Within USAID and the larger international population community, evaluation is often discussed in terms of performance versus impact evaluation. Although these two types of evaluation differ in many respects, one important difference pertains to data sources. In most cases, the evaluation of performance (to use the terms cited above) will require program–based data. By contrast, evaluations conducted to measure outcome (e.g., contraceptive prevalence as an intermediate outcome or fertility as a long–term outcome) require population–based data. Figure I–3 summarizes the relationship between the types of indicators (input – process – output – outcome) and the sources of data (Reynolds, 1990).

Figure I–3

Levels of Indicators in Family Planning Program Evaluation

Program–based (Performance)	Population–based (Outcome)
■ Input	■ Effect (Intermediate)
■ Process	■ Impact (Long–term)
■ Output	

The reason for opting for these terms is to avoid the possible confusion surrounding the word "impact," which through widespread non–technical usage no longer retains its precise meaning in terms of evaluation. For example, an I–E–C specialist might speak in terms of the impact of an ongoing media campaign on new acceptors in local clinics. This case is very different from a U.S. Congressman who wants to know the impact of USAID's family planning programs (on fertility).

The distinction between program–based[5] and population–based is as follows. Program–based data consist of information available from program sources (e.g., administrative records, client records, service statistics) or information that can be obtained from on–site collection (e.g., observation of provider–client interaction, mystery client surveys). Also, a follow–up study of clients who had attended a clinic would use program–based data, in that the names of the clients would come from the clinic records.

Within the program level it is important to further differentiate the components. Inputs (program resources) are fed into processes (program activities), which in turn produce output (program results) and ultimately outcome (population–based results), as shown in the following sequence:

Input → Process → Output → Outcome

Inputs are human and financial resources, physical facilities, equipment, and operational policies that enable services to be delivered.

Process refers to the multiple activities that are carried out to achieve the objectives of the program. It includes both what is done and how well it is done.

Although a high level of input is generally reflected in a satisfactory process of service delivery, it is theoretically possible to have a high level of input but a poorly delivered service (for example, if a high level administrator opposed to family planning were successful in blocking service delivery in facilities under his/her control). Conversely, there are countless real–life examples around the world where service providers with highly inadequate resources nonetheless strive to provide the best service they can under the circumstances.

Output refers to the results of these efforts at the program level. Although family planning program managers at the field level are interested in national trends regarding prevalence and fertility, they will tend to limit the evaluation of their own activities to program–based measures, especially measures of output. Two types of output, shown in separate boxes in Figure I–1, are service output (that measure the adequacy of the family planning service delivery system) and service utilization (that measures the extent to which the services are used).

By contrast, the evaluation of outcome refers to measuring the effect that the program has on the larger social system, usually the general population of a given target area (e.g., the population of a specific country). It can also refer to a smaller area (e.g., the catchment area for a demonstration project), provided that the data are drawn from a random sample of that population. The Demographic and Health Survey is a primary source for population–based evaluations of family planning program activities.

Within the category of population–based evaluation, it is important to distinguish between two kinds of outcome: intermediate and ultimate (long–term).

Effect (intermediate outcome):

- that which is a relatively direct and immediate result of program process and output (e.g., contraceptive prevalence).

Impact (ultimate outcome):

- that which is an anticipated result of program process and output in the long–term (e.g., change in fertility rates), but also subject to the influence of many non–program factors (such as socio–economic conditions or status of women).

Many family planning programs are designed to reduce fertility; however, it often takes years to produce this impact. Moreover, if and when it is achieved, it is not always possible to make a causal link between the family planning program in question and the fertility decline (given that other factors such as increased levels of education or improved economic conditions may also contribute to the outcome).

Thus, program evaluations often concentrate on intermediate level outcome (effects), which are seen as more directly linked to program effort and which are expected to reflect change in a shorter

[5] An alternative term for program–based is "facility–based." However, this term implies some type of service delivery point, which might not be the case for certain indicators collected by the program (e.g., number of persons trained in I–E–C). Thus, the term program–based is generally used in the Handbook.

period of time. The most widely reported intermediate outcome of family planning programs is contraceptive prevalence, the percentage of women of reproductive age currently using a contraceptive method.

The maturity of the family planning program will determine in part the type of evaluation strategy to use. For example, all programs, fledgling or mature, can benefit from the monitoring of process and impact. However, in programs that are at the emergent or launch stages (Destler et al. 1990), it is appropriate to focus largely on process and output rather than population–based outcome. Even highly successful programs require a number of years to show an increase in contraceptive prevalence at the national level. Those eager to demonstrate outcome a year after the program is launched may, in fact, invite frustration and discouragement over the apparent lack of effect. By contrast, mature programs still benefit from fine tuning that results from the monitoring of process and output; yet the results of primary interest will relate to outcome.

It should be noted that in the conceptual framework, all of the boxes in Figure I–2, up to and including the service output, lend themselves to performance evaluation.[6] The output from the different functional areas (e.g., a strategic plan to orient family planning program effort, trained personnel applying new skills, a functioning logistics system in place, I–E–C messages that inform potential clients of the location of services, and so forth) contribute collectively to defining the service output, based on accessibility, quality of services, and program image/acceptability.

In contrast, the boxes at the far right on the conceptual framework (Figure I–1) reflect the intended outcome of family planning programs, which call for population–based indicators (e.g., contraceptive practice, fertility rates, and other measures of health/social status).

Implicit in this conceptual framework are feedback loops. The results obtained on output indicators may require changes in program input; the results obtained for service output may require reexamination of the activities undertaken (processes) in different functional areas.

One important caveat is that these indicators are descriptive in nature; when collected over time, they provide important trend data (e.g., increases in the number of functioning service delivery points, percentage of target audience exposed to family planning messages, number of couple–years of protection, level of contraceptive prevalence, and so forth). However, they do not indicate causation. Even if service delivery is improving (as measured by indicators listed in the Handbook) and contraceptive prevalence increases over time, one cannot conclude that improved service delivery is responsible for the increase in prevalence, based on the descriptive data alone. Methodologies are available for determining causal relations (such as experimental designs and longitudinal multi–level analysis); in fact, these methodologies typically use the types of indicators listed in the Handbook as dependent variables. However, in the absence of such designs or analytic techniques, these indicators do not allow for a determination of cause–and–effect.

In concluding this discussion on types of indicators, it is important to note that this Handbook deals largely with objective rather than subjective or qualitative indicators. Within the social science research community, there is a growing awareness of the importance of both types of information in applied research and evaluation. Qualitative indicators are useful primarily in process evaluation. The Handbook does include some subjective indicators (e.g., in the section on quality of care); yet the individual evaluators would be well advised to develop this aspect of the evaluation approach more fully.

Sources of Data

The indicators described in this Handbook require multiple sources of data. The type of data and nature of the institutions responsible for generating the data are summarized in Figure I–4. These institutions include:

- government offices and institutions,
- independent organizations (universities, research firms, management consultants), and
- national family planning programs.

---

[6] The one exception relates to the Information–Education–Communication (I–E–C). The key indicators used to determine whether I–E–C activities are on target include both measures of output and of effect at the population level (e.g., percent of target population exposed to family planning messages).

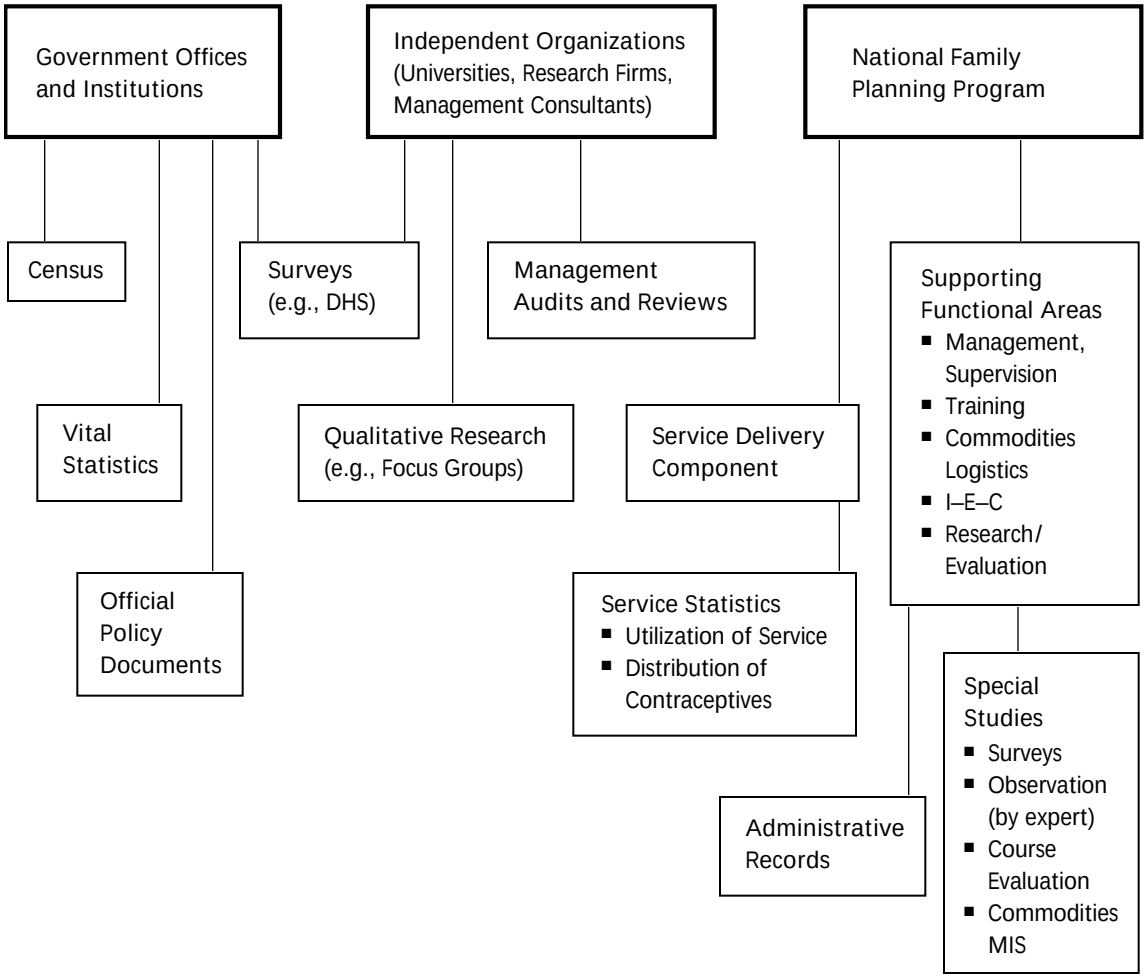
Over a lifetime, a family planning program might well use all or most of the sources and types of data mentioned in Figure I–4. However, a specific evaluation exercise is more likely to draw on these data sources one at a time. For example, country specialists interested in program issues would analyze the Demographic and Health Survey (DHS) for all indicators related to family planning (e.g., demand for services, reach of I–E–C programming, contraceptive use, and so forth).[7] Program managers interested in monitoring service utilization might exploit whatever indicators were available from the program's management information system. Lest the reader be discouraged by the sheer number of boxes in Figure I–4, it should be stressed that they represent the full menu, only certain items of which are selected by a given institution for a specific evaluation exercise.

Programs operating under budgetary constraints (which is the vast majority) tend to use sources of data that can be generated at minimal cost (e.g., service statistics, administrative records)

[7] Many of the indicators in this Handbook are based on data from the Demographic and Health Survey (DHS) or similar national surveys on fertility and reproductive health. For a description of the DHS–type survey, see the Introduction to Chapter VIII.

Figure I–4

Sources and Types of Data for Family Planning Program Evaluation



or to access data for which the cost is borne by others (e.g., the DHS). Where evaluation questions cannot be answered from existing sources of data (for example, assessing quality of care in a given program), special studies are required.

In Chapters II–IX, the description of the indicators includes the most common source(s) of data for each. In addition, the indicators available from the different sources of data are summarized in Appendix A.[8]

This Handbook does not include instructions on how to collect the data for each indicator. One useful reference in this regard is Chapter 8 on "Collecting Data" in *Improving Family Planning Evaluation. A Step–by–step Guide for Managers and Evaluators* (Garcia–Nuñez, 1992). Some of the topics are also covered in standard texts on data collection for social science research. For data collection in a specialized area (e.g., a commodities and logistics management information system), one may find it useful to consult experts in the area.

Scoring of Indicators

The indicators in this Handbook include dichotomous, ordinal, and interval measurements. The dichotomous measurements generally constitute indicators that would be scored as "yes" or "no" (e.g., the existence of a formal population policy). Ordinal measures include indicators for which ratings such as excellent, good, fair, or poor would be appropriate, though a specific numerical value might not be (e.g., quality of program leadership). The continuous measurements include rates, percentages, scores on knowledge tests, and so forth (ratio of public sector family planning expenditures to GDP).

It would be beneficial for the field of family planning evaluation to develop standardized indices, particularly with respect to components of the family planning supply environment (i.e., policy environment, service delivery operations, and service output). This would allow for an assessment of progress in these areas for a given country over time and possibly for cross–national comparisons. Although there are limitations to such standardization, the model developed by Lapham and Mauldin in the form of the Family Planning Program Effort Score illustrates the utility of such indices (Lapham and Mauldin, 1985).

Work is in progress regarding scoring mechanisms. The Population Council is in the process of developing scoring rules for use in connection with the Situation Analysis (Mensch and Jain, 1991). The Commodities and Logistics Working Group has developed an index on the "state of the commodities/logistics system" that incorporates numerous ordinal measurements that collectively yield a quantitative score (see Chapter III).

At this time, the development of standardized indicators to measure different aspects of the family planning supply environment is a goal rather than a reality. Such indices take time to develop; moreover, the indicators to be included in the indices must be tested in field situations to determine the feasibility and validity of this approach. For this reason, this edition of the Handbook does not include scoring rules to be used in connection with the indicators. However this topic will be revisited in the second edition of the Handbook.

We encourage any colleagues working on the development of such scoring rules or standardized indices in specific field applications to provide us with information pertaining to their experiences or recommendations regarding useful approaches to this task.

Anticipated Update of this Handbook

The EVALUATION Project began in October 1991 and will continue through September 1996. This document constitutes the first edition of this Handbook, which has been prepared early in the project in an effort to obtain consistency in the

[8] The major source of data for population–based indicators is the DHS survey. A new interactive software package, "EASEVAL," has been developed under The EVALUATION Project in conjunction with DHS staff, to facilitate the use of the DHS data from standard recode files. The package is currently available for use with DHS I files; it will be available for use with DHS II once the standard recode files for DHS II are released. The menu for this software allows the user to select "INDICATORS," which gives a complete listing of the indicators described in this Handbook that are available from the DHS survey. In appendix A, all indicators available from the DHS interactive software package are marked with a superscript[1] to highlight the link between this Handbook of Indicators and the EASEVAL package.

indicators and definitions used in connection with project activities.

However, as the indicators presented in this Handbook are subjected to further empirical testing, it will be useful to update this document. Also, it is expected that as working groups are convened under The EVALUATION Project on additional topics (information–education–communication, and evaluation), new indicators will come to light. Accordingly, it is currently anticipated that the Handbook will be updated in 1995 (the fourth year of the project). Users of this edition are encouraged to contact the project staff with comments and suggestions for useful modifications in the second edition.

With this as background, we proceed to define commonly used indicators in the evaluation of international family planning programs.

- Existence of a policy development plan
- Number of appropriately disseminated policy analyses
- Number of awareness–raising events targeted to leaders
- Existence of a strategic plan for expanding the national family planning program
- Integration of demographic data into development planning
- Number of statements of leaders in support of family planning
- Formal population policy addressing fertility and family planning
- National family planning coordination
- Level of the family planning program within the government administration
- Levels of import duties and other taxes
- Restrictions on advertising of contraceptives in the mass media
- Absence of unwarranted restrictions on providers
- Absence of unwarranted restrictions on users
- Public sector resources devoted to family planning as a percentage of GDP
- Quality of program leadership
- Extent of commercial sector participation

INDICATORS TO MEASURE THE POLICY ENVIRONMENT

The policy environment of family planning programs is defined as the factors affecting program performance that are beyond the complete control of national program managers. In addition to political support and other expressions of national policy (e.g., a formal national population policy), the policy environment includes those aspects of operational policy which involve decisions at a higher level than the program (i.e., the program's organizational structure, its legal/regulatory environment, the resources made available to it and its use of provider and acceptor payments and fees).

Figure II–1 on page 19 provides a conceptual framework for the evaluation of the policy area. The framework is organized according to the standard Input–Process–Output–Outcome schema and depicts policy activities of a single period as part of a continuous circular loop. The policy environment is the output of the policy process. It directly affects the various functional areas of family planning programs (e.g., I–E–C, training, commodities and logistics, management), institutionalization, self–sufficiency and family planning demand.[1] The current chapter presents a set of process and output indicators for the policy area.

Inputs to the policy development process include:

- the external environment;
- domestic policy inputs; and
- donor inputs.

The external environment includes a country's political–administrative system (PAS), its socio–economic characteristics, its socio–cultural environment, and its family law. Domestic policy inputs include available data, existing research, staff resources of policy units, equipment (e.g., computers, audio–visual equipment), and domestic funding. Domestic inputs are enhanced over time to the extent that the institutionalization of policy development capabilities is an effect of policy work (although Figure II–1, as a single–period schema, does not explicitly show the feedback effect from institutionalization in one period to levels of domestic policy inputs in the following period). Donor inputs to policy development include specialized technical expertise, equipment, funding, international research, policy dialogue, non–project assistance and conditions precedent to loans and grants. Although no input indicators are presented in the current chapter, any evaluation of policy development activities should confirm that the needed inputs have been delivered as planned.

The policy environment is modified over time through the planned implementation of policy activities (i.e., the process of policy planning and policy development). Policy development planning is based on an assessment of the current policy environment in relation to program needs and of the inputs available for further policy development. Many policy development activities, or policy interventions, are designed to strengthen political support and/or to develop an effective national policy in support of family planning programs. Increasingly, however, as support for family planning grows at the national level, policy interventions are being directed to strengthening the operational policy environment. Five indicators of the process of policy planning and policy development are presented in this chapter:

Policy Planning

- Existence of a policy development plan

Policy Development

- Number of appropriately disseminated policy analyses

[1] A compete description of the linkages depicted in Figure II–1 can be found in Knowles, Bollen and Yount (1993).

- Number of awareness–raising events targeted to leaders
- Existence of a strategic plan for expanding the national family planning program
- Integration of demographic data into development planning

As shown in Figure II–1, the external environment (directly), other policy inputs (indirectly), and the process of policy development determine a national family planning program's policy environment.[2] The dimensions of the program policy environment, which is the output of the policy development process, include:

- political support;
- national policy; and
- operational policy.

Political support, at the national, regional and local levels, plays a central role in a program's policy environment since it is an important determinant of the other dimensions of the policy environment. Political support can be both explicit and implicit. *Statements made by high–level government officials and other leaders in support of family planning* is a commonly used indicator of explicit political support. Implicit political support is most often gauged by what the government actually does in the areas of national and operational policies.

National policy includes both formal statements of policy (e.g., national population policies, national development plans) and tax and other material incentives designed to affect parents' fertility decisions.[3] Operational policy consists of four sub–dimensions which are directly related to the operation of national family planning programs:

- *organizational structure and processes*: a family planning program's status within the government's administrative structure and its capacity to mobilize the resources of other public and private institutions;
- *the legal/regulatory environment*: taxes and other restrictions that affect the supply of contraceptives, particularly from the private sector, and medical barriers to family planning service delivery;
- *provision of resources*: the financial, material, and human resources needed by family planning programs; and
- *pricing*: fees charged to clients (cost recovery) and payments to providers and acceptors (incentives).[4]

Eleven output indicators referring to the program policy environment are presented in this chapter:

Political Support

- Number of statements of leaders in support of family planning

National Policy

- Formal population policy addressing fertility and family planning

Organizational Structure and Processes (Operational Policy)

- National family planning program coordination

---

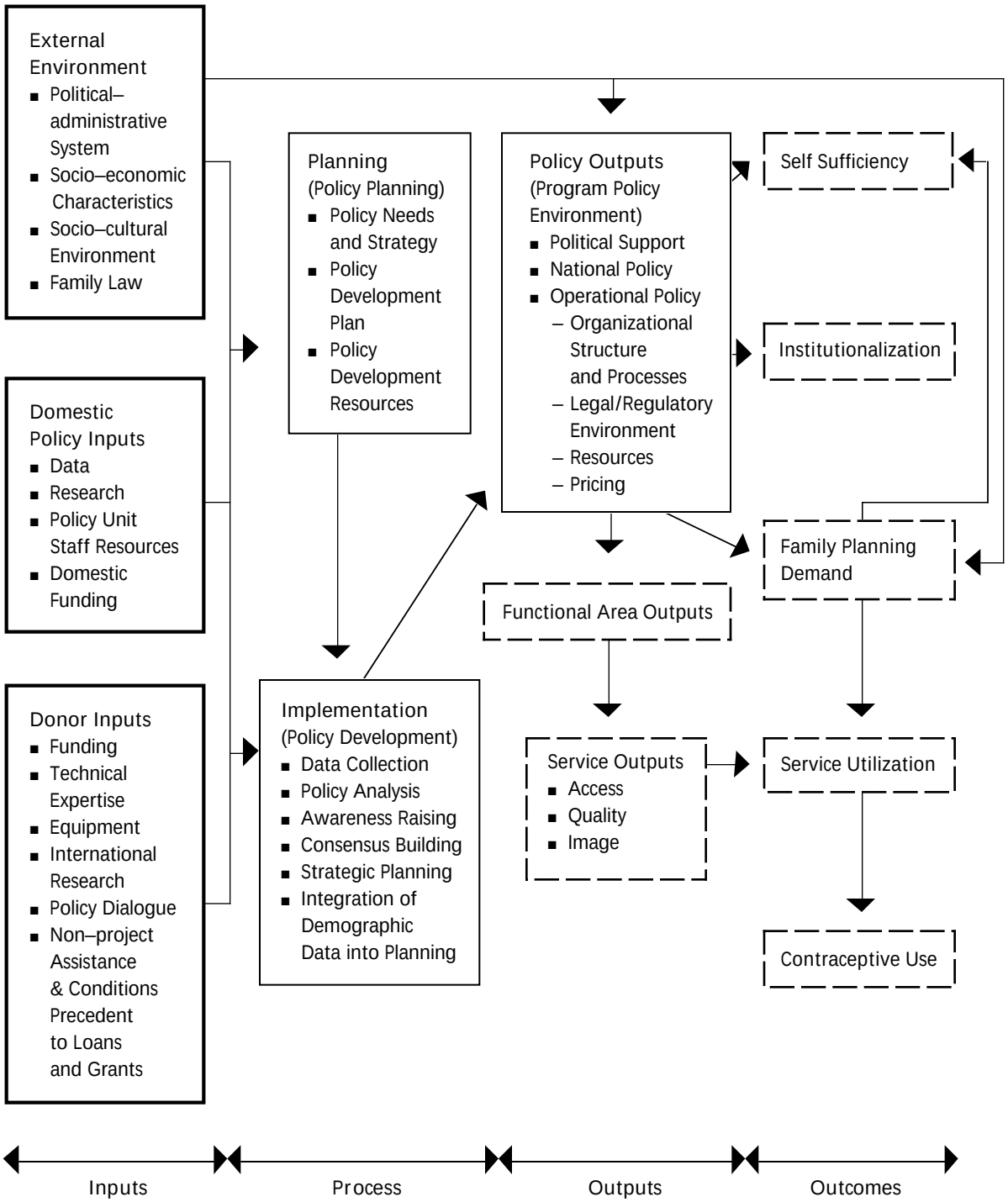
[2] Although the external environment is, in most cases, beyond the control of a family planning program and its leaders, it is important to keep in mind the dominant role it may play in affecting the policy environment. Perhaps to a greater extent than with other functional areas of a family planning program, all the appropriate activities may be carried out to produce a favorable policy environment but the output may not be attained due to unfavorable developments in the broader external policy environment.

[3] National policy incentives are limited to those expressly designed to affect fertility and family size but which are not linked directly to contraceptive use. Incentives related to the actual use of contraceptives (i.e., provider and acceptor incentives) are considered in the present framework to be part of operational policy. Family law intended to promote other social goals but which has secondary effects on fertility (e.g., minimum legal age at marriage, inheritance law) is considered to be part of the external environment.

[4] Although the use of incentives is often indicative of strong political support for family planning, it is also controversial, particularly when it appears to encourage use of certain methods (e.g., sterilization) or involves coercion. Incentives are less controversial when they are used to overcome constraints to family planning use (e.g., to reimburse client travel costs or to improve provider performance). Distinguishing between these cases, however, is often difficult in practice. No indicator referring to incentives is proposed in this chapter for general use.

Figure II–1

Conceptual Framework for Evaluation of Policy



- Level of the family planning program within the government administration

Legal/Regulatory Environment
(Operational Policy)

- Levels of import duties and other taxes
- Restrictions on advertising of contraceptives in the mass media
- Absence of unwarranted restrictions on providers
- Absence of unwarranted restrictions on users

Resources
(Operational Policy)

- Public sector resources devoted to family planning as a percentage of GDP
- Quality of program leadership
- Extent of commercial sector participation

According to Figure II–1, improvements in the program policy environment should lead to stronger service delivery (access, quality, image), increased service utilization and contraceptive use, as well as to enhanced institutionalization and self–sufficiency of programs. As noted above, institutionalization also affects levels of domestic policy inputs in the following period (a feedback loop). On the supply side, therefore, the policy environment contributes directly both to improved service delivery in the short run, and to enhanced program sustainability in the long run. On the demand side, both the political support and national policy dimensions of the program policy environment (e.g., statements of leaders, incentives) affect family planning demand.

It is important to note that many of the policy indicators proposed in the Handbook have equivalents (or near equivalents) among the Lapham/Mauldin/Ross Family Planning Program Effort measures, currently available for 98 countries for 1982 and 1989 (Lapham and Mauldin, 1985; Ross et al. 1988, 1992). These data can serve as valuable baseline data for the evaluation of policy development activities. The relationships, if any, between the indicators proposed here and corresponding Lapham/Mauldin/Ross indicators are discussed below in the descriptions of the individual indicators.

Indicator

EXISTENCE OF A POLICY DEVELOPMENT PLAN

Definition

This is a qualitative (yes/no) indicator. A "yes" value is assigned if a policy development plan exists at the national level which: (1) assesses the current policy environment for family planning; (2) identifies and provides a strategy for removing important policy obstacles; (3) assigns responsibility for policy development activities identified in the strategy; and (4) provides a budget adequate to ensure the implementation of the strategy.

Data Requirements

The policy implementation plan, which may be part of a larger planning document.

Data Source(s)

The unit within the national population program which is responsible for policy development.

Purpose and Issues

This process indicator of policy planning activities measures progress made in developing a plan which views policy development as a systematic process designed to achieve specific policy objectives. In the absence of such a plan, policy development activities may be random in purpose, poorly sequenced and not targeted to appropriate audiences.

A policy development plan should be developed for all programs and should be revised over time to reflect changes in the policy environment or in program strategies and needs. In some cases, a policy development plan may call for the development of a formal policy, if one does not already exist. If a formal policy exists, it is particularly important that the plan provide for its implementation.

Indicator

Number of Appropriately Disseminated Policy Analyses

Definition

Policy analyses are designed to provide actionable information to policy makers to encourage them to improve one or more aspects of the policy environment of family planning programs. Examples of possible analyses include: (1) analyses of the development impact of rapid population growth (e.g., RAPID analyses[5]); (2) analyses of the demand for family planning services; (3) market segmentation analyses; (4) legal and regulatory analyses; (5) target–cost analyses; (6) cost–benefit analyses; and (7) cost–effectiveness analyses. "Appropriately disseminated" implies that each analysis is designed to address an important policy obstacle and is targeted to a suitable audience and that its findings have been disseminated through an appropriate channel and format to this target audience. "Number" refers to a given period (e.g., one year).

Data Requirements

Description of each policy analysis, including its policy objectives, the target audience and a description of the manner in which the completed study was actually disseminated.

Data Source(s)

Administrative records of those organizations carrying out the various studies.

Purpose and Issues

The purpose of this process indicator is to provide a quantitative measure of one type of activity undertaken as part of the process of policy development.

This is a simple measure of activity that in no way reflects either the quality of the effort or its impact on policy output. It is useful to the extent that it creates a sense of accountability among staff responsible for these activities.

It is important to distinguish between this indicator and the following indicator (*Number of Awareness–raising Events Targeted to Leaders*). The present indicator is designed to reflect the number of technical studies prepared and disseminated, whereas the following indicator is designed to reflect the number of events. A single policy development activity (e.g., a RAPID subproject) may be reflected in both indicators if an analysis is prepared which is then disseminated through a number of awareness–raising events. On the other hand, the results of studies may be disseminated through other channels; and awareness–raising events may not involve studies (e.g., conferences).

[5] RAPID is the acronym for the A.I.D. funded Resources for the Awareness of Population in Development project, which prepares computer–assisted presentations for policymakers on the development impact of lower fertility and strong family planning programs.

Indicator

Number of Awareness–raising Events Targeted to Leaders

Definition

"Events" may include conferences, workshops, RAPID presentations, fairs, media campaigns, and observational travel[6] designed to make decision–makers in health and other sectors more knowl–edgeable about and conscious of population/family planning issues. "Number" refers to a given period (e.g., one year).

Data Requirements

Number of events, listed by type of activity, numbers and official positions/responsibilities of persons attending or participating.

Data Source(s)

Administrative records of those organizing these activities.

Purpose and Issues

The purpose of this indicator is to provide a quantitative measure of a commonly used policy intervention. The proposed indicator is a simple measure of activity that in no way reflects either the quality of the effort or its impact on policy output. It is useful to the extent it creates a sense of accountability among staff responsible for these activities.

---

[6] Observational travel refers to trips which are arranged for governmental officials and other persons in key positions to visit and observe successful population/family planning programs in other developing countries (usually in the same region). Similarly, the purpose may be to observe activities to improve the policy environment, such as the steps involved in developing a formal population policy.

Indicator

Existence of a Strategic Plan for Expanding the National Family Planning Program

Definition

This is a qualitative (yes/no) indicator. A "yes" value is assigned if there exists a long–term plan at the national level which: (1) defines the objectives of the national family planning program over a 5– to 10–year period, including quantitative targets; (2) defines a clear strategy for attaining these objectives; (3) establishes an organizational structure for the program which is consistent with the strategy, covering both the public and private sectors; and (4) projects the resources (material, human and financial) required to implement the strategy, as well as sets forth a plan to secure them.

Data Requirements

An approved long–term plan for the national family planning program.

Data Source(s)

The government organization responsible for coordinating the national family planning program.

Purpose and Issues

This is a process indicator of an important policy development activity. Its purpose is to measure whether a national family planning program has developed a clear view of its mission and objectives and of its strategy for attaining them. Strategic planning at the national level requires the participation of a variety of government ministries, including typically, those of Health, Finance, Planning, Information, Education, and Interior, as well as important private groups (commercial and NGOs , religious organizations, women's groups).

A strategic plan at the national level should address such issues as: (1) the emphasis to be placed on *demand creation* versus the provision of public services; (2) the appropriateness of the current method mix; (3) the respective roles of the public, NGO and commercial sectors; (4) the manner in which the program will be financed; and (5) the role of various ministries in carrying out the plan.

Strategic planning at the national level should be paralleled by similar planning at the level of each institution which is part of the overall program. However, strategic planning at the institutional level is an aspect of sound management; at the national level, strategic planning is a policy exercise involving the participation of all major actors.

Indicator

Integration of Demographic Data into Development Planning

Definition

This is a qualitative (yes/no) indicator. A value of "yes" is assigned if long–term plans exist, both for the overall economy and for individual sectors (e.g., health, education, food, housing, water, environment) which show clearly how population growth impacts on the attainment of development objectives.

Data Requirements

Five–year development plans; perspective plans covering a longer period (e.g., 10 years); sectoral plans.

Data Source(s)

Ministry of Planning; National Population Council (or equivalent population policy coordinating body); line ministries (e.g., Health, Education, Agriculture, Housing, Water, Environment).

Purpose and Issues

This indicator measures the extent to which government planners and other mid–level technocrats understand population dynamics and their impact on development and have incorporated this understanding into their plans. Both A.I.D. and UNFPA have provided assistance in recent years to population planning units to encourage the integration of demographic data into development planning. This process indicator attempts to measure the degree of success experienced by this type of policy intervention.

Increased integration of demographic data into development planning was one of the recommendations of the 1974 World Population Conference in Bucharest. Donors have assisted this effort in many countries in the hopes that planners would become strong advocates of fertility reduction and family planning. Although advocacy on the part of planners has undoubtedly occurred in some countries, there is little systematic evidence that it has led to stronger political support for family planning. Such a result is not surprising, however, given that the main preoccupation of planners is better planning–not fertility reduction or family planning.

Indicator

Number of Statements of Leaders in Support of Family Planning

Definition

"Statements" may be delivered in public speeches or appear in written form in official communications/documents (e.g., national development plans). "Leaders" refers primarily to high–level government officials who are responsible for governing the country, establishing its laws, and allocating its public funds to the different sectors. However, it may also extend to other types of leaders (e.g., religious, medical) who, for whatever reason, speak with authority about population and family planning. "Number" refers to a given period (e.g., one year).

Data Requirements

Evidence of such statements, including dates and sources, position/responsibilities of the person making them, intended audience and media employed.

Data Source(s)

Texts (or video tape) of official speeches, newspaper articles, government communiques, official documents, other public expressions.

Purpose and Issues

The purpose of this indicator is to measure the degree of explicit political support for family planning on the part of key government officials and other leaders.

Many would consider this a fairly "soft" indicator, since it is difficult to quantify and interpret. Even if one is able to assiduously follow all official speeches and documents of numerous high level decision–makers, it is unclear how to assess such statements. For example, how many such statements are "enough?" One a year? One a month? Moreover, a single statement by the president of the country might carry more weight than 100 statements by lower level officials. Nonetheless, this indicator is included because it may be one of the few available markers of progress in the initial stages of policy work. Also, such statements may constitute the earliest signals of impending change in the government's position on family planning.

The impact of such policy statements is greater if they continue over a long period of time and if much the same commitment to family planning is made by successive leaders. If, on the other hand, successive leaders vacillate between strong support and weak or no support, the policy impact of such statements may be minimal.

This indicator is similar to the Lapham/Mauldin/Ross Family Planning Program Effort measure, *Favorable statements by leaders*: "Whether the head of government speaks publicly and favorably about family planning at least once or twice a year, and whether other high officials also do so."

Indicator

Formal Population Policy Addressing Fertility and Family Planning

Definition

This is a qualitative (yes/no) indicator. A "yes" value is assigned if a formal population policy has been officially adopted, disseminated, and implemented which addresses the problem(s) of high fertility and/or closely spaced births and which endorses family planning as an important step to address such problems. A formal population policy, which is either a written document or part of one, is an official statement of a government that establishes goals and (in some cases) targets for the population sector and includes a strategy for attaining them.

Data Requirements

Text of any formal policy; evidence of official acceptance of the policy; reference to the national population policy in speeches of high level officials; reports of dissemination and implementation of the policy.

Data Source(s)

Public laws and official government documents; newspaper articles, government communiques, other public expressions; organization responsible for implementing the policy.

Purpose and Issues

This output indicator is a measure of the degree to which national policy supports fertility decline and family planning.

The adoption of a formal population policy is often seen as a key accomplishment in the area of policy. It is important to keep in mind, however, that what counts most is its dissemination and implementation. It is also noted that many successful family planning programs have been conducted in the absence of a formal policy.

A formal policy that is developed from domestic initiative, rather than from external initiative or pressure, is more likely to be broadly disseminated and implemented and, therefore, to achieve significant impact. The policy impact of such a formal statement will also depend on how much emphasis the document gives to family planning and whether it sets quantitative targets for contraceptive use.

This indicator is similar to the Lapham/Mauldin/Ross Family Planning Program Effort measure, *Government's official policy or position concerning fertility/family planning and rates of population growth*: "Existence and type of official policy to reduce the population growth rate, to support family planning activities in the absence of government–sponsored activity, or to discourage family planning services." The United Nations Population Division has also conducted seven periodic surveys of member countries since 1977 about their official policies and positions on population growth, fertility levels and family planning. These data can also serve as a valuable source of baseline data for the evaluation of policy development activities in individual countries.

Indicator

NATIONAL FAMILY PLANNING COORDINATION

Definition

This is a qualitative (yes/no) indicator. A value of "yes" is assigned if there is a functioning body (i.e., one which meets at least once each year) establishing policy for family planning and coordinating both public and private sector family planning service delivery. Examples include national population councils and national family planning boards. Public sector membership of such bodies is typically at the level of minister or vice minister, with the chairman often the prime minister or minister of planning; private sector membership typically includes representatives of various PVOs, organizations of private medical providers and religious groups. A permanent secretariat often exists, along with an interministerial working group composed of senior technocrats from key ministries.

Data Requirements

Statutes establishing such bodies (e.g., national population policies); minutes and reports of meetings; reports of any secretariat established to provide backup to such a coordinating body.

Data Source(s)

Legal codes; administrative decrees; text of national population policies; ministry or organization housing; the secretariat of the national coordinating body.

Purpose and Issues

This output indicator is intended to measure whether there is effective coordination within the national family planning program between various ministries (e.g., health, education, planning) and between the public and private sector organizations delivering family planning services. It is a measure of the organizational structure (operational policy) dimension of the program policy environment.

It is important to recognize that some countries have nominal coordinating bodies which either do not meet (or meet rarely) or which play only a marginal role. In such cases, this indicator should be given a "no" value. It should also be given a "no" value in cases where a functioning body exists but works against, rather than in support of, the family planning program.

Finally, it is noted that a number of countries have well–coordinated national programs even in the absence of a formal coordinating body. The existence of a formal, interministerial coordinating body is not, therefore, a necessary condition for a strong policy environment.

Indicator

Level of the Family Planning Program Within the Government Administration

Definition

This indicator measures the position of the family planning program director in the government administration (i.e., number of administrative levels down from the head of state to which the program director reports).

Data Requirements

Government organizational charts.

Data Source(s)

Ministry in which family planning program is situated.

Purpose and Issues

This output indicator measures the organizational strength of a family planning program within the government administration. It is a measure of the organizational structure (operational policy) dimension of the program policy environment. This indicator is similar to the Lapham/Mauldin/Ross Family Planning Program Effort measure, *Level of family planning program leadership*: "Level of the post (person appointed) to direct the national government family planning program, and whether or not the program director reports to the highest level of government."

Indicator

LEVELS OF IMPORT DUTIES AND OTHER TAXES

Definition

This indicator is defined as the percent of the retail price of contraceptives that is attributable to import duties and other taxes, whether they are produced locally (e.g., import duties on imported raw materials, value–added taxes) or imported directly.

Data Requirements

Data on import duties and domestic tax rates.

Data Source(s)

Government tax codes and customs schedules.

Purpose and Issues

The purpose of this output indicator is to measure the extent to which various taxes affect the retail price of contraceptives. It is a measure of economic barriers (operational policy) to the supply of contraceptives from the commercial sector.

It is important to note that the percentage impact on retail prices of taxes levied on producer costs may be substantially increased if wholesale and retail margins are applied to these taxes. For example, if producer taxes on each oral contraceptive cycle increase producer prices from $0.20 to $0.40 per cycle and if the retail price of $1.00 is determined by adding a fixed wholesale and retail margin of 150 percent to the producer price, the above indicator should be given a value of 50 percent, rather than 20 percent (since the effect of the tax is to double the retail price).

This indicator is related to the Lapham/Mauldin/Ross Family Planning Program Effort measure, *Import laws and legal regulations regarding contraceptives*: "Extent to which import laws and legal regulations facilitate the importation of contraceptive supplies that are not manufactured locally, or the extent to which contraceptives are manufactured within the country."

Indicator

Restrictions on Advertising of Contraceptives in the Mass Media

Definition

This indicator is qualitative, with four levels: (1) no restrictions; (2) brand–level advertising permitted only for non–prescription products (e.g., generic advertising of oral contraceptives is allowed, but not that of specific brands); (3) no advertising permitted for prescription products; and (4) no advertising of any contraceptives permitted.

Data Requirements

Regulations and restrictions on mass media advertising.

Data Source(s)

Ministry of Information, reports of social marketing projects and other projects working with the commercial sector.

Purpose and Issues

This output indicator is a measure of legal/regulatory barriers (operational policy) to the supply of contraceptives, particularly from the commercial sector (i.e., the inability to advertise contraceptives restricts potential market demand, discouraging the commercial sector from entering the market).

This indicator is similar to the Lapham/Mauldin/Ross Family Planning Program Effort measure, *Advertising of contraceptives in the mass media allowed*: "Whether the advertising of contraceptives in the mass media is allowed with no restrictions, whether there are weak restrictions, whether there are social restrictions, or whether there are strong restrictions."

Indicator

Absence of Unwarranted Restrictions on Providers

Definition

This indicator has a maximum value of five points. One point is given for each of the following conditions: (1) appropriately qualified and trained paramedical personnel are permitted to prescribe oral contraceptives, administer injections and insert IUDs; (2) appropriately trained community–based distribution (CBD) workers are allowed to resupply oral contraceptives; (3) pharmacies are permitted to dispense oral contraceptives without a prescription and to administer injectables (if permitted to administer other types of injections); (4) properly trained and equipped general practitioners are permitted to perform sterilizations and to insert and remove implants; and (5) properly trained and equipped physicians are permitted to perform abortions on demand.

Data Requirements

Medical regulations.

Data Source(s)

Ministry of Health; legal codes.

Purpose and Issues

This output indicator measures the extent to which medical barriers (operational policy) impose restrictions on providers of family planning services. Although the rationale for and against medical barriers is often couched in technical terms, medical barriers are a policy matter to the extent that their existence reflects underlying political forces. Additionally, decisions to remove them may entail high–level policy decisions.

This indicator is defined on the basis of a set of medical barriers which are important in many countries. The list of barriers may need to be altered somewhat in individual country applications. An important criterion to apply in selecting the most relevant medical barriers for inclusion in the indicator is their expected impact on actual provider behavior. For example, restrictions on CBD workers will have no impact where there is no CBD program. In other cases, restrictions may be on the books but not followed in practice. Interviewing providers may yield valuable information as to which medical barriers have the greatest impact on current practices.

Indicator

Absence of Unwarranted Restrictions on Users

Definition

This indicator has a maximum value of four points. One point is given for each of the following conditions: (1) no medically unwarranted restrictions (e.g., age, sex, marital status, number of surviving children, gender of surviving children, spousal consent, unwarranted medical contraindications) on the use of any contraceptive are imposed on acceptors; (2) no medically unnecessary tests are required of acceptors or of continuing users; (3) continuing users of oral contraceptives are given at least three months' resupply with each visit; and (4) abortion is legal and openly available.

Data Requirements

Medical regulations and clinical practices.

Data Source(s)

Ministry of Health regulations; legal codes; situation analyses and other user/provider surveys; service statistics.

Purpose and Issues

This output indicator measures the extent to which medical barriers (operational policy) impose restrictions on clients. As noted with respect to the previous indicator, the existence of medical barriers may reflect underlying political forces; and their removal may involve high–level policy decisions. Also, as with the preceding indicator, the list of specific barriers included in the indicator may vary in individual country applications. Individual medical barriers should be included in the definition of the indicator only in cases where they have demonstrable impact on user behavior. Interviews with users and providers may yield reliable information for use in selecting the most important medical barriers to incorporate in individual country applications.

Indicator

Public Sector Resources Devoted to Family Planning as a Percentage of GDP

Definition

This indicator is defined as the ratio of total public sector recurrent expenditures on family planning to gross domestic product (GDP), expressed as a percentage (usually less than one percent).

Data Requirements

Public sector recurrent expenditures on family planning in current prices; national accounts data on GDP in current prices.

Data Source(s)

National recurrent expenditure budgets; national accounts.

Purpose and Issues

This output indicator is a measure of the commitment of resources (operational policy) by a host country government to its family planning program. To the extent that donors finance any of the family planning program's recurrent expenditures (most donor support is to the capital budget), this indicator may not be an accurate reflection of host country government effort.

It is important that the indicator reflect whether funds allocated in budgets are actually expended to provide family planning services. Many governments fall short of implementing their published budgets. It is also important to confirm that the resources expended on family planning programs are actually used for this purpose and that the resources flow to the operational units in the field providing services.

Particularly in countries where family planning services are provided along with other MCH services, there may be no readily identifiable line item in the appropriate ministry's/organization's recurrent budget that can be linked to family planning. Moreover, where personnel are engaged in providing other health services in addition to family planning, it may be difficult to allocate some proportion of their time to family planning. In such cases, one of several alternative procedures can be employed. The simplest and most commonly used approach (but probably the least reliable) is to interview supervisors and health workers, asking them to estimate the percentage of their time spent providing family planning services. This percentage can then be used as a basis for allocating labor and other joint costs. Alternatively, a time use survey can be carried out in a sample of facilities, using either the technique of patient flow analysis or direct observation of health workers at specified intervals (i.e., work sampling).

This indicator is similar to the Lapham/Mauldin/Ross Family Planning Program Effort measure, *In–country budget for program*: "Percentage of the total family planning/population budget available from in–country sources. A top score is given if in–country sources provide 85% or more of the budget; no score is given if these sources provide less than 50% of the budget." This indicator is difficult to compute, however, if it includes (as it should) private sector financing as part of the domestic share. It is also subject to exchange rate fluctuations since the denominator includes donor contributions expressed in U.S. Dollars (or other form of foreign exchange).

Another commonly used indicator of government resource commitment to family planning is the *Share of the national budget allocated to family planning.* The main problem with this alternative indicator is that family planning programs are often financed by several levels of government (e.g., national, state, local). Another problem is that such an indicator would be sensitive to variations in the size of the national budget due to political, ideological or national security considerations. The indicator proposed above avoids these limitations by including public sector expenditures by all levels of government and by expressing them relative to a measure of the country's overall level of economic activity (GDP).

Indicator

Quality of Program Leadership

Definition

This indicator should be based on an index of the performance of program directors as managers, covering such dimensions of performance as experience, continuity, and commitment, technical skills, ability to motivate other program managers and staff, ability to mobilize domestic resources and support from other ministries, ability to communicate program achievements and needs to the public, ability to work within the bureaucracy, ability to work effectively with donors, and ability to plan strategically.

Data Requirements

Systematic evaluation of the performance of the family planning program director and other key program officials (e.g., Minister of Health, head of national population commission), including time spent in position.

Data Source(s)

Independent evaluations of external observers.

Purpose and Issues

This output indicator reflects the quality of the human resources (operational policy) committed by a government to its family planning program. Some observers consider it to be one of the most important indicators of implicit political support.

This indicator should be limited to the one, two or three (at most) officials who have the greatest influence on program directions. In many cases, the program director alone will suffice. Knowledgeable persons (e.g., donor staff, NGO staff) can be interviewed to rate the individual involved on the basis of criteria such as those listed above.

This indicator may be extremely revealing as to the commitment of the government to population/family planning. Programs cannot be expected to flourish in the absence of dedicated and sustained leadership.

One problem with this indicator is that the ratings on which it is based are necessarily subjective and may therefore be influenced by program outcome (e.g., whether the program has been successful in raising contraceptive prevalence levels).

In addition to the criteria listed above, there is an important subjective dimension relating to the attitudes of persons in key leadership positions toward family planning. There are cases where persons appointed to head a population commission have been openly hostile to family planning. Thus, to the extent possible, this too should figure into any evaluation of the "qualifications" of key personnel.

There is no related indicator in the Lapham/Mauldin/Ross Family Planning Program Effort measures. However, the policy development literature cites quality of program leadership as one of the most important factors in the policy environment affecting program success or failure (Finkle and Ness, 1985; Ickis, 1987; Lapham and Simmons, 1987).

Indicator

EXTENT OF COMMERCIAL SECTOR PARTICIPATION

Definition

This indicator is defined as the percentage of modern–method contraceptive prevalence that is accounted for by the commercial sector.

Data Requirements

Contraceptive prevalence, by method and source.

Data Source(s)

DHS or other contraceptive survey.

Purpose and Issues

This output indicator measures one aspect of the resources (operational policy) dimension of the program policy environment: the involvement of the commercial sector (e.g., private providers, pharmacies) in family planning service delivery. This is an indicator of the program's organizational maturity and of its success in mobilizing private sector resources to meet growing demand for services. In most societies, as programs mature, the commercial sector's share of prevalence grows relative to both the public and NGO sector shares. This development is likely to occur due to efforts made by the public program to promote sustainability (e.g., introduction of user fees for those able to pay for services) and due to the growing demand on the part of the public for high quality and convenient family planning services.

It is important to note that some programs may opt to remain largely public sector programs, due either to the weakness of the commercial sector or to strong support for (and sufficient resources to fund) a largely public sector program. Examples of mature programs continuing to rely on the public sector are those of China and Thailand.

In some cases, this indicator may be high simply because the government program is weak. Since the indicator would give a misleading impression of the policy environment in these circumstances, it should probably not be used in such cases.

Lastly, it is noted that the size of the commercial sector depends on socio–economic factors (e.g., household income, urbanization) as well as on government policy. In very poor, predominantly rural areas, a weak commercial sector may not be a valid indicator of the policy environment.

The related Lapham/Mauldin/Ross Family Planning Program Effort measure, *Involvement of private–sector agencies and groups*, reflects only the involvement of the NGO sector: "Extent to which private–sector agencies and groups assist with family planning or other population activities. These groups include family planning associations, special service groups (e.g., for sterilization services), religious associations, and so on. The involvement or assistance with family planning and population activities may include the following: delivery of family planning supplies and services, training, family planning information and education, membership in a family planning interagency group that meets at least twice annually, moral support, or other types of assistance."

Chapter III

Service Delivery Operations

A. Management

B. Training

C. Commodities and Logistics

D. Information–Education–Communication (I–E–C)

E. Research and Evaluation

Indicators for each functional area are listed at the beginning of the corresponding sections.

INDICATORS TO MEASURE SERVICE DELIVERY OPERATIONS

Service delivery operations refer to the different components or operations within a family planning program: management,training, commodities/logistics, information–education–communication (I–E–C), and research/evaluation. Also known as functional areas or subsystems, these operations correspond closely to the divisions found in family planning organizations worldwide. Also, they constitute an important link in the family planning supply environment, as shown in Figure I–2 of Chapter I.

The family planning supply environment has been a relatively neglected area of family planning evaluation. Historically, the primary focus has been on intermediate– (e.g., contraceptive prevalence) and long–term outcome (e.g., fertility decline). Less attention has been devoted to the factors that precede these outcome in the conceptual framework, namely the supply–side factors that contribute to the demand for family planning services.

As noted in Chapter I, two broad categories of factors influence the demand for family planning services. On one hand, a series of social, economic, religious, and cultural factors influence societal norms and individual preferences with regard to family size. As demonstrated by Lapham and Mauldin (1985) and confirmed by Mauldin and Ross (1991), fertility decline is associated with socioeconomic well–being of the population. Conversely, lack of socio–economic development with its related social consequences has worked to sustain a demand for large families in many countries.

However, the strength of existing family planning programs also plays a role in fertility decline. In these same two articles cited directly above, fertility decline was higher in those countries that scored high on the family planning program effort index. The family planning supply environment can counter some of the pressures for sustained child–bearing and can influence the demand for children, first by offering services to those who already feel the need to space births or limit family size, and second, by creating an awareness of the alternatives to continued child–bearing among others.

The strength of the different functional areas collectively determines the adequacy and performance of the family planning service delivery system in a given country, measured as service output. The three dimensions of service output are access, quality, and image (illustrated in Figure I–2 of Chapter I and presented in detail in Chapter IV). It is not hard to visualize the arrows between any of the service operations described in this chapter and the elements of access, quality, and image, outlined in the next. For example, training increases the number of service sites able to provide contraceptives and the quality of the services given. I–E–C informs the population of their access to services, promotes the quality of those services, and thus enhances the image of the program.

Those working in one of the specific functional areas of service delivery often express the desire to demonstrate the results of their area on what "really counts" in family planning: outcome at the population level. Given the attention focused on demonstrating the demographic impact of family planning in recent years, experts in these functional areas may have felt pressure to "prove" that what they do has a direct payoff in terms of contraceptive prevalence or even fertility decline.

Scientifically, the optimal approach to demonstrating the impact of a given functional area is a controlled field experiment. In practice, such studies are usually done on a small scale basis (often as an operations research project), not for the country as a whole. In the context of a national family planning program, it would be difficult to hold other components of the program constant in

order to assess the contribution of a specific functional area, especially in light of some resistance at the field level to consciously withholding a potentially beneficial service or communication from any part of the population. Consequently, assessment of the unique contribution of a specific functional area to the overall program effect is methodologically possible but almost never done.[1]

It is useful to conceptually trace the contribution of a specific functional area to the effect of the program at the population level (as is done in the first chapter in Figures I–1 and I–2). However, it is not possible to analyze what percentage of change (e.g., in contraceptive prevalence) can be attributed to a given functional area, and much less, what percentage of a fertility decline could be attributed to that functional area. Technically speaking, a full system of indicators can allow for the study of each area's contribution to the total demographic picture. However, the primary use of these indicators is to monitor what the areas provide vis–à–vis total program effort.

Most evaluations of the functional areas within family planning programs are not designed to "prove" or quantify their contribution to the overall effort, but rather to identify areas where improvements are warranted, which would lead to a more effective program.

In this chapter, we present indicators for five functional areas, listed in the same order in which they are presented in the conceptual framework (Chapter I, Figure I–2). In the case of Management, Training, and Commodities/Logistics, the indicators described below were developed in the context of working groups on these topics, organized under The EVALUATION Project. While an Evaluation Working Group has yet to be convened, the indicators presented in the section on Research and Evaluation benefited from the discussions of the Operations Research Working Group during its two meetings to date. By contrast, the I–E–C section constitutes a draft, to be further developed in the context of future working group meetings on that topic.

---

[1] The issue of evaluating the unique contribution of a specific functional area to program outcome should not be confused with the feasibility and desirability of evaluating the effects of a program in its entirety (reflecting the contribution of all the functional areas). In the latter case, a controlled field experiment is the method of choice, although to date this has been greatly underutilized. Programs have often missed valuable opportunities to conduct "natural experiments," the design of which builds on the siting (i.e., strategic geographical placement) of interventions in selected areas and the timing of these interventions.

MANAGEMENT

Illustrative Indicators:

- Existence of a clear mission that contributes to the achievement of program goals
- Realization of operational targets
- Clearly defined organizational structure
- Adequacy of staffing
- Awareness of current financial position
- Access to current information on key areas of program functioning
- Access to current information on program progress
- Capacity to track commodities

Management is a term that is often loosely defined and, as a result, poorly understood. This reflects, in part at least, the tendency for perceptions of what management is and should be to vary across cultures and organizations (Laurent, 1983; Hofstede, 1982; Finkle and Ness, 1985). Despite this, it is more or less universally acknowledged that a great deal of the credit for the success of family planning programs in developing country settings is due to program management.

The materials presented in this section represent the initial product of efforts by a Management Working Group convened under The EVALUATION Project to consider the vast quantity of literature on and field experience in the management of family planning programs in order to develop a conceptual framework and specify a set of indicators for this functional area. It is anticipated that the proposed indicators for the functional area will be closely scrutinized, tested and refined over the course of The EVALUATION Project. The results of further development efforts, as well as feedback from persons and organizations working in the management area, will be reflected in the update to this Handbook.

The starting point in considering indicators for the functional area is to establish a working definition of program management. As indicated in the introduction to this Handbook, a family planning program is defined as an organized program designed to provide information, supplies, and service of modern means of fertility control to those interested. National family planning programs often consist of a national coordinating body and a dominant service delivery organization, complemented by smaller organizations that serve specific program functions and/or market segments. The lead service delivery organization is often, but not always, an organization affiliated with the government. In fact, public– and private–sector entities have effectively cooperated in a variety of different configurations to provide family planning services in different parts of the world (Ickis, 1987).

Management has been defined alternately in terms of "what management is" or "what management does." Warwick (1988), for example, defines management as the design, organization, implementation, evaluation and redesign of family planning programs and operations. For Finkle and Ness (1985), management is defined in terms of

the primary tasks that managers undertake. They define the principal task of management as motivating and directing humans so that the resources that they represent are directed toward the achievement of organizational goals. In either case, management is seen as involving direct influence over the allocation of work and rewards within the organization, and is thus distinguished from coordination, which implies only partial or indirect influence. Management is a concept that applies to the national family planning program as an entity, as well as to each of the contributing organizations.

Irrespective of definitions, program management is viewed as having primary responsibility for, given a set of input and constraints, the production or generation of program–level output and population–level outcome. It is the role of management to guide service delivery operations toward the production of intended output. This is accomplished through three primary types of activities or processes: planning, implementation, and evaluation.

In considering indicators for the management functional area, it is useful to think of effective management as consisting of a series of essential attributes or characteristics, which may be viewed as underlying dimensions or elements. Eight key dimensions or elements may be identified:

- mission,
- planning/strategy,
- organization,
- human resources,
- finance,
- information,
- monitoring and evaluation, and
- logistics.

A preliminary list of 63 indicators was developed for the eight elements by the Management Working Group. Subsequent refinement resulted in the list of indicators shown in Appendix B of this Handbook. Note that in many cases it was deemed useful to organize the indicators under a particular element in terms of sub–elements, as shown in Appendix B.

Due to space limitations, it is not possible to discuss each of the proposed indicators in detail in this Handbook. Instead, one "illustrative" indicator for each element was chosen for discussion. The indicators chosen for presentation were those that were thought to pertain most closely to "the bottom line" for a given element; that is, to reflect the end product or output of management efforts along a particular dimension. This approach is consistent with the view of management as a results–oriented process. It is also consistent with the "rule of contingency" proposed by Finkle and Ness (1985), which holds that there is no one best way of organizing anything: what matters is the bottom line.

Accordingly, it is anticipated that in using the indicators proposed in the Handbook, the starting point for assessment of any given element would be the "bottom line" indicator. If program management is performing well on this indicator for a particular element, it is likely to be performing well on the other aspects measured by more detailed or specific indicators for the element. If performance is found to be lacking on the bottom line indicator, the other indicators measuring specific factors or aspects that contribute to the bottom line result may prove useful in identifying areas or aspects that require further attention.

As a final note, it should be recognized that in some instances the "boundaries" between the management functional area and other functional areas of service delivery discussed in this Handbook may not be readily apparent. This is due to the fact that a management component spans and is thus involved in each of these areas. In other words, the success (or lack thereof) in performance in any of the functional areas is at least in part attributable to management. For example, program performance in the Training or I–E–C areas is to some extent determined by management's recognition of its importance and by the resources allocated to its implementation.

However, to the extent that the influences of effective management on program output or outcome are often greater than the sum of the individual activities undertaken, it was felt that the examination of indicators of overall management would contribute additional insights into program functioning. This point should be borne in mind when reviewing the indicators presented in this section. In actual application, indicators from both the section on general management and the more specific functional areas might be considered in assessing a particular functional area.

Indicator

Existence of a Clear Mission That Contributes to the Achievement of Program Goals

Element

Mission

Definition

The program has a written statement of mission that, should it be achieved, would make a significant contribution to the achievement of national reproductive health and family planning goals.

A mission is considered to contribute to the achievement of national family planning goals if it:

- is consistent with national policies and priorities regarding family planning and reproductive health;
- provides a vision of the future (for, at minimum, 5–10 years);
- defines the program's services and products;
- defines the program's target clients; and
- is sufficiently clear and detailed that it provides a meaningful basis for developing operational plans.

Data Requirements

Evidence of a written mission statement; information on the program's mission; information on national reproductive health and family planning policies and goals.

Data Source(s)

Program documents.

Purposes and Issues

This indicator provides a summary measure of the existence of a clear sense of direction in the form of a written mission statement. Although it is possible for programs without formally–defined missions to make important contributions to the achievement of national reproductive health/family planning goals, a clear mission serves the important function of helping to keep program management and staff focused on the accomplishment of long–term objectives. A written mission statement, often included in strategic planning documents, communicates the sense of mission to internal and external audiences. As noted above, the mission statement should contain a number of key elements that provide the basis for the development of operational plans.

A mission or policy statement that defines goals that are adapted to the internal and external realities of the program (i.e., is appropriate and realistic within the context in which the program operates) would receive the highest score on this indicator. A program lacking a formal statement would receive the lowest score. The clarity and coherence of the statement and its appropriateness for the program at its current state would determine scoring between the two extremes.

Indicator

Realization of Operational Targets

Element

Planning/Strategy

Definition

Program operational planning targets over a defined reference period (e.g., the past 2–3 years) have been successfully met.

Data Requirements

Information on operational planning targets for the period covered by the assessment; evidence on the extent to which targets have been reached.

Data Source(s)

Program planning documents; performance reviews and other output of monitoring and evaluation efforts.

Purposes and Issues

This indicator provides a "results–oriented" measure of program operational planning performance, that is, the extent to which planning targets are realized. As a "bottom line" indicator, it presupposes the existence of defined strategic and operational plans with measurable objectives and targets. Programs that do not undertake strategic and operational planning cannot be meaningfully evaluated (and will thus score poorly) on this indicator, since they will not ordinarily have clearly specified and measurable targets and objectives.

However, the proposed indicator goes beyond the mere existence of strategic and operational plans to try to get at the issue of whether the plans are meaningfully utilized to guide service delivery operations. By focusing on the end results of strategic and operational planning, the indicator is also likely to provide at least a partial measure of the extent to which strategic and operational plans are realistic given program resources and existing constraints.

The indicator also presupposes the existence of a program monitoring and evaluation system that provides information on progress made toward the accomplishment of targets and objectives on a regular basis (see element 7 below for further discussion of monitoring and evaluation systems as a key element of management). Programs without means of measuring performance against objectives will not be rated highly on the indicator.

Finally, it should be noted that the validity of the indicator rests on the assumption that strategies and plans serve to focus program resources and activities in ways that enhance program effectiveness. Given this, demonstrated capacity to develop and realize plans is viewed as an attribute of effective management.

Indicator

CLEARLY DEFINED ORGANIZATIONAL STRUCTURE

Element

Organization

Definition

The program/organization has an organizational chart and accompanying document(s) that define the relationships among organizational units, the managerial chain of command, and the roles and responsibilities of staff at all levels of the organization as the program actually operates.

Data Requirements

Evidence that program unit interrelationships, the management chain of command, and roles and responsibilities of staff have been formally defined and accurately describe how management decisions are made in the organization.

Data Source(s)

Program documents and interviews with program staff.

Purposes and Issues

This indicator is an illustrative indicator of one aspect of organization structure. The reader is referred to Appendix B for a listing of other sub-elements and indicators under the element of organization proposed by the Management Working Group.

The indicator is intended to provide a measure of the extent to which roles and responsibilities within the program are clearly defined, both "on paper" and in practice. Note that the indicator does not advocate any particular organizational structure. Rather, it attempts to measure the extent to which, whatever the structure of a program, it is sufficiently well defined and that managers and staff know their roles and responsibilities within the organization.

The validity of the indicator is based on the assumption that a clear definition of the chain of command and position roles and responsibilities in an organization or program helps to ensure that strategic and operational decisions are made so as to make optimal use of available resources. It is also assumed that the program structure is (at least partially) amenable to change directed by program management, and is thus a valid criterion on which to evaluate management functioning.

Assessment of program structure and roles and responsibilities entails an analysis of relationships between units and staff within the program. Such an assessment would normally begin with an analysis of the program organizational chart and job descriptions of program staff. An organizational chart that clearly defines boundaries and bridges between the different functional units and individuals within the program would receive the highest score for this aspect of structure. By contrast, a program with undefined relationships, the extreme being the absence of an organizational chart, would receive the lowest score.

A second consideration is the extent to which program decision–making processes conform to those embodied in the formally defined program structure. This involves assessing the appropriateness of staff for the positions they fill and the extent to which they actually make the decisions called for in the (theoretical) organizational scheme. The latter might be accomplished, for example, by reviewing how one or more recent important decisions were made. It is assumed that a program with a sound organizational scheme that has been put into practice will be more effective in achieving its objectives than one that exists on paper alone. For example, some programs have organizational schemes that call for decentralized decision–making (on at least routine matters) but in actual practice decision–making tends to be highly centralized. The result is often an organization that is slow to respond to opportunities to improve performance and one with high levels of staff hesitancy to make decisions that they may well be in the best position to make.

Assessment of the latter aspect will tend to be more subjective than the former. One approach might be to use a scale ranging from an exact or close conformity between the theoretical structure of decision–making at one extreme to little or no conformity at the other.

Again, as in other "bottom line" indicators in this functional area, programs that score poorly on this indicator would benefit from scrutiny on other, more specific indicators of program structure as a means of identifying areas for improvement.

Indicator

ADEQUACY OF STAFFING

Element

Human resources

Definition

All staff positions identified in the program staffing plan are (except for a small number of temporary vacancies) filled by personnel who have the qualifications and competencies required for the position as stated in the position description.

"Competency" refers to the fact that staff have sufficient skills to carry out the functions or activities called for by a given position.

Data Requirements

Information on qualifications and skill requirements (by position); information on qualifications and competence of staff occupying each position.

Data Source(s)

Program personnel documents and records; personnel management information system data; staff competency assessments.

Purposes and Issues

This indicator provides an overall measure of management performance in the area of personnel management. The indicator measures actual performance against a "gold standard," the scenario in which all positions are filled by staff who fully satisfy the training and competency requirements indicated for the position in the program personnel structure.

The indicator assumes the existence of: (1) skills–based position descriptions indicating the qualifications and competencies required for each position; and (2) a personnel management information system component or sub–system that tracks current personnel assignments and competencies of staff filling each position. In programs that lack either of these components, special staff competency assessments will have to be undertaken in order to measure this indicator.

The measurement of competency is discussed in greater detail in the presentation of indicators of Training in the next section of the Handbook.

Indicator

AWARENESS OF CURRENT FINANCIAL POSITION

Element

Finance

Definition

Management can provide current information on amounts budgeted and expended for major budget line items.

Data Requirements

Information on the level of knowledge of program managers on amounts budgeted and expended for the current year, by major line item.

Data Source(s)

Interviews with program manager; program budget documents and financial statements.

Purposes and Issues

This indicator provides a summary measure of program financial management capacity. It equates capacity to budget and track expenditures for major categories in the budget with adequate financial management. The indicator assumes the existence of (1) a strategic plan and (2) a financial management system and supporting information sub–system that enables managers to track revenues and expenditures. Programs lacking these cannot be meaningfully evaluated on the indicator, and it is thus highly likely that improvements in financial management would be needed.

An alternative "bottom line" indicator of financial management capacity might be "the extent to which revenues sufficient to implement the strategic plan are in place or may be reasonably expected." This indicator provides a measure of both the program's capacity to take advantage of opportunities to raise revenues, and the extent to which strategic plans are realistic given the program's capacity (and "track record") to generate revenues.

Indicator

ACCESS TO CURRENT INFORMATION ON KEY AREAS OF PROGRAM FUNCTIONING

Element

Information

Definition

Program management can produce up–to–date information on key areas of program functioning, consisting of:

- facilities and equipment,
- personnel,
- commodities and logistics,
- finance, and
- service statistics.

See below for definitions of minimum information requirements for each of these sub–systems.

Data Requirements

Evidence of the availability of current information covering the areas outlined above; for example, routine reports from a management information system (MIS), annual or periodic program reviews or reports that are based on information generated by the program, or the routine use of information generated by the program in undertaking performance reviews and annual workplan development.

Data Source(s)

Program planning and performance assessment documents; review of management information system protocols and output; interviews with program managers and supervisors.

Purposes and Issues

This indicator provides a "bottom line" measure of the adequacy of information support to program management. The focus of the indicator is on the availability of current information, which is viewed as essential to effective management.

The indicator assumes the existence of a management information system (MIS) or sub–systems that support(s) planning and decision–making in each of the key areas of program functioning outlined above. Programs that do not have an established MIS will generally not have current information available to management. Partial "scores" are possible for this indicator, since some programs may have reasonably current information available for some areas (e.g., service statistics and commodities and logistics) but not others. Programs that have more fully–developed management information systems will score higher on this indicator than programs with only partially developed information support systems.

In scoring the indicator, the following might be considered the minimum information content of each sub–system.

- Facilities and equipment: number, location, and services provided at all service delivery points; annual inventory of equipment, including location and current condition
- Personnel: total staff, by training/qualifications and current assignment; number of vacant positions
- Commodities and logistics: quantities procured (by commodity) during the current year, quantities in stock at central stores, quantities disbursed from central stores during the current year, authorized inventory levels at service delivery/supply points, year to date number of stock–outs at service delivery/supply points
- Finance: amounts budgeted and year to date expenditures, by major budget line item
- Service statistics: number of family planning service visits, number of new acceptors, method mix of new acceptors

It should be noted, however, that the indicator does not take into account how effectively program managers utilize the information available to them. Assessment of this dimension would require the consideration of yet more "bottom line" indicators; for example, measures of program effectiveness and efficiency. See Section E of this chapter for a discussion of several such indicators.

Indicator

ACCESS TO CURRENT INFORMATION ON PROGRAM PROGRESS

Element

Monitoring and evaluation

Definition

At any given point in time, program managers and staff can produce reasonably up–to–date data describing progress toward the accomplishments of the targets, objectives and goals that have been established in program operational and strategic plans.

The definition of "reasonably up–to–date" varies by type of information. See below for further discussion of this point.

Data Requirements

Evidence of the availability of relevant information on progress made toward accomplishing program targets and objectives derived from monitoring and evaluation activities.

Data Source(s)

Program documents and program review/evaluation reports.

Purposes and Issues

This indicator provides a summary measure of the extent to which program management and staff has access to information that enables them to track progress toward the accomplishment of program objectives and goals. The indicator assumes the existence of targets and objectives that have been set through strategic and operational planning exercises. The indicator also assumes the existence of a management information system that adequately supports management decision–making (see the previous indicator in this section) and program capacity to undertake research and evaluation activities as needed (see Section E of this chapter for discussion of indicators in this area). Programs that do not have these capabilities will not score well on this indicator.

Program staff having access to information on accomplishments and areas requiring further attention is deemed an important aspect of the present indicator, since this would require the existence of "feedback" mechanisms from top program management, a factor which has been associated with improved program performance (Finkle and Ness, 1985).

The inclusion of the qualifier "reasonably up–to–date" in the definition of the indicator is meant to highlight the need for monitoring and evaluation activities that are scheduled and undertaken in a manner that is consistent with the time frame in which important program management decisions are made. For example, to the extent that process evaluation (or monitoring) activities are to be used to inform annual program reviews or workplan development, information for previous years would normally be required. Data pertaining to performance against annual workplan targets that are not available for several years after the fact are not considered to be reasonably up–to–date.

Other types of evaluation results (for example, operations research findings, effectiveness or efficiency evaluation results, etc.) are not required as frequently, but for programs to score highly on the present indicator, they need to be available in advance of the time when they are to be used (for example, in developing a new five–year strategic plan). Indicators of program capacity in the area of research and evaluation are discussed in greater detail in Section E of this chapter.

Indicator

CAPACITY TO TRACK COMMODITIES

Element

Logistics

Definition

Program staff responsible for logistics management are able to provide an up–to–date account of the following:

- quantities procured (by commodity) during the current year,
- quantities in stock at central stores,
- quantities distributed from central stores during the current year,
- authorized inventory levels at service delivery/supply points, and
- number of stock outs, by location and type of service/supply point, during the current year.

Data Requirements

Evidence that the logistics management data described in the definition of the indictor above are available to program managers.

Data Source(s)

Program records; commodities and logistics management information system.

Purposes and Issues

This indicator provides a summary measure of program management's capacity to track and thus effectively manage commodities procurement and distribution. Adequate information support is assumed to be a prerequisite for effective commodities and logistics management. The indicator thus assumes the existence of: (1) a commodities and logistics operational plan, and (2) a logistics management information system (LMIS). Managers in programs that do not have a LMIS will not be able to provide the information required for the indicator.

It should be noted that this indicator is not a "bottom line" indicator in the same sense of the indicators proposed for other elements in this functional area. More "bottom line" indicators would be, for example, "the frequency of stock outs" or "the proportion of service delivery points stocked according to plan." However, since commodities and logistics has been identified as a key service delivery function in its own right (see Section C in this chapter for a discussion of indicators specific to this functional area), the emphasis of commodities and logistics–related indicators under the heading of management is more on the existence of underlying management systems. Again, as noted in the introduction to the present chapter, the "boundaries" between some of the functional areas identified in the Handbook are in some cases arbitrary and, accordingly, users of the Handbook may choose to use indicators from different chapters/sections as they are deemed relevant to the evaluation of a particular program activity.

TRAINING

- Number/percentage of courses that achieve learning objectives
- Number/percentage of courses that contribute to the achievement of program training objectives
- Number/percentage of courses where the training methodology is appropriate for the transfer of skills and knowledge
- Number of trainees by type
- Number/percentage of trainees who have mastered relevant knowledge
- Number/percentage of trainees competent to provide a specific family planning service
- Number/percentage of trained providers assessed to be competent at a specified period (e.g., 6 months) post–training
- Number/percentage of trainees who apply the skills to their subsequent work

This section draws heavily on the work of the Training Working Group organized under The EVALUATION Project. Much of the momentum for this activity came from an earlier Evaluation Working Group (convened by the Information and Training Division of the Office of Population).

Training as used in this chapter refers to different types of events, ranging in duration from one day to several months, which have the specific purpose of increasing knowledge and developing/improving the skills of participants to perform specific tasks. In this version of the Handbook, we focus on indicators that would be most useful in training for clinic–based delivery of contraceptive services. Analogous indicators are needed for training in other domains (e.g., in any one of the functional areas listed in the table of contents).

Training—as well as the other functional areas—contributes to the objective of increasing prev–alence rates, reducing fertility, and improving maternal and child health status. However, it has this effect by strengthening the family planning supply environment: in increasing access to services, improving the quality of the services (e.g., better counseling, improved clinical skills of providers), and enhancing the image of the program. The indicators in this section measure both the process of carrying out training activities and output with respect to the supply environment.

The output of training can be measured at the level of the individual trainee (or "participant") and at the level of the family planning service delivery system. For example, one indicator is the number/percentage of trained providers that apply the skills to their subsequent work. Its analogue at the system level is the number of service sites where trained providers are regularly providing services. Since the system–level indicators overlap with the indicators used to assess the family planning supply environment (e.g., access in Chapter IV), this chapter will focus primarily on the individual level indicators (i.e., the level of the trainee).

The training indicators that follow are the product of several meetings of the Training Working Group. The minutes from these meetings give additional detail on "system–level" indicators that are analogous to the individual level indicators

described here. While the final report of the Training Working Group (Bertrand and Brown, 1993) lists all indicators developed in the series of meetings, this Handbook includes those that are more widely used by training organizations at the present time.

Indicator

Number/Percentage of Courses That Achieve Learning Objectives

Definition

"Objectives" are those outlined in the course curriculum or syllabus.

Data Requirements

[If assessed by participants] Response to the question, "In your opinion, did the course meet the objectives outlined in the first session?"

[If assessed by independent observer with expertise in the content area] Review of the course content and observation of trainees' skills.

Data Source(s)

Course evaluation by participants upon its completion; or notes of independent course observer.

Purpose and Issues

The purpose of this indicator is to determine whether the course content provides trainees with the knowledge and skills outlined in the course objectives. Course evaluations are widely used in training sessions for service personnel. Observation by an independent observer with expertise on the topic is more common in training of trainer courses.

Course evaluations are subject to a courtesy bias, especially if participants question the confidentiality of the exercise or if they have developed a positive interpersonal relationship with the trainers over the course of the event. This bias can best be reduced by stressing the confidentiality of the evaluation form.

Indicator

Number/Percentage of Courses That Contribute to the Achievement of Program Training Objectives

Definition

"Program training objectives" are ideally the result of an institutional or national needs assessment. This indicator is measured by the number/percentage of training events that contribute to a specific training objective and the profile of participants in each.

Data Requirements

Inventory of program training objectives, number of events designed to contribute to each of the training objectives, and the profile of participants for each event.

Data Source(s)

Administrative records (Training Division).

Purpose and Issues

This indicator answers the question: to what extent did the program succeed in completing the quantity and type of training events outlined in its training plan? It is generally used in relation to a specific objective; the key information is the number of courses given by type (e.g., the number of courses that contributed to the objective of training providers of IUDs, the number of courses that contributed to the objective of training master trainers). It does not measure the adequacy of the training events, per se.

The absolute number of training events (by type of training) must be interpreted in relation to the number judged necessary by the experts who developed the training plan. Thus, if only one training of trainers course was deemed necessary, and it was conducted with the anticipated number/type of participants, then "1" would constitute satisfactory performance on this indicator. By contrast, the same program might have conducted five refresher courses for CBD workers; but if this represented only half of the number stipulated in the training plan, then performance would be assessed accordingly.

This indicator assumes that the training plan was developed based on an adequate needs assessment of the local situation. If the contrary were true, then the results for this indicator would be of questionable value.

Indicator

Number/Percentage of Courses where Training Methodology is Appropriate for the Transfer of Skills and Knowledge

Definition

"Training methodology" refers to the series of techniques and approaches used in the event; "skills/knowledge" are specific to each course or event.

Data Requirements

List of all training events; objectives for each event; and assessment tools specific to that event.

Data Source(s)

Course evaluation forms completed by participants or notes of an expert observer hired to assess participants' skills.

Purpose and Issues

This indicator is used by training organizations for internal evaluation of their own activities. For example, if a trainer were hired to conduct a clinical IUD skills course, but only gave lectures, this would be considered an inappropriate training methodology for transferring skills.

Observer comments are routinely collected by training organizations as part of project documentation. Often project directors or others associated with the training will note this in their reports back to the organization.

This indicator is reported in reference to specific training courses, but is not generally aggregated over courses, since its main purpose is to provide feedback for improving the training methodology in specific settings or on specific topics. However, as part of an external evaluation, this indicator could be reported as an aggregate measure.

Indicator

NUMBER OF TRAINEES BY TYPE

Definition

"Trainee" refers to a participant in any type of training event, regardless of its duration; "type" refers to the different categories of participants (e.g., physicians, nurses, social workers) or the subject matter covered (e.g., IUD insertion, NORPLANT® implant insertion, voluntary surgical contraception).

Data Requirements

Number of persons (based on an actual list of names for potential verification purposes).

Data Source(s)

Records, usually kept by the Training Division, which are used both for administrative purposes during the training (e.g., distributing per diem) and for monitoring of trainees at a later date.

Purpose and Issues

The indicator serves as a crude measure of activity, and it can be used in determining whether a program/project meets its target and/or in tracking progress from one year to the next. However, the "unit of measurement" is not strictly speaking uniform, in that one "trainee" may have attended a course for one day, whereas another participated for three months.

The measure can be further improved in several ways:

- by expressing the number of trainees by category of training;
- by expressing the number of trainees as the percentage of the number scheduled for training in a given year;
- by expressing the number of trainees as a percentage of the estimated number needed to be trained to fill a national program mandate, determined through a systematic needs assessment prior to the initiation of training activities; and
- by expressing the number successfully completing the course as a percentage of the total that attend the course.

Indicator

NUMBER/PERCENTAGE OF TRAINEES WHO HAVE MASTERED RELEVANT KNOWLEDGE

Definition

"Mastered relevant knowledge" must be operationally defined in terms specific to a given context. "Mastery" is conventionally used in relation to acquisition of knowledge. ("Competency" involves both knowledge and skills; see next indicator.)

Data Requirements

Listing of individuals; evidence of mastery of knowledge.

Data Source(s)

Administrative records (training files); written test (e.g., pre– and post–tests of knowledge).

Purpose and Issues

This commonly used indicator in the evaluation of training measures the participants' ability to retain key information in the short term. Low scores reflect inadequacies in the course and/or the inability of participants to absorb the information. Every training organization that has developed or uses training manuals has identified the knowledge that should be acquired by a category of trainees on a specific subject. This knowledge is tested by pre– and post–tests.

The test results indicate whether the trainee understands certain key points, even though the number and definition of key points will differ by context. The items included in the test should be those most relevant to a particular training exercise. If the same questions are used on subsequent tests, this indicator can be used to monitor trends over time within a program.

One limitation with respect to this indicator is a lack of standardization regarding the items on these tests. Some training organizations have a list of questions that they encourage host country organizations to adopt for testing purposes, but some countries opt to design their own questions. This lack of standardization makes it difficult to compare the results from this indicator across countries and even across programs within a given country. A second problem is a lack of standardization regarding the word "mastered." For example, in some countries, a passing grade could be 60%, whereas in others the required score for passing would be 100%.

Despite these limitations, training organizations routinely use this indicator as a means to control the quality of training conducted in connection with their activities. For example, course certification may be dependent on a passing grade, which has meaning in a particular country; and in some countries, providers must have the certificate before being allowed to deliver services.

Indicator

Number/Percentage of Trainees Competent to Provide a Specific Family Planning Service

Definition

"Competent" refers to the fact that the trainee can deliver the service according to a set standard, which may differ according to the training context. Thus, the evaluator must know the standard of the context. "Competent" is used by training organizations in reference to the acquisition of skills; however, since knowledge is generally a necessary part of acquiring a skill, the term competency covers both knowledge and skills.

Data Requirements

Assessment of trainees against established standards for a number of family planning tasks, conducted by an expert observer.

Data Source(s)

Checklist of the expert observer.

Purpose and Issues

This indicator serves to measure the technical competence of participants at the completion of training or thereafter with respect to specific skills. It reflects both the adequacy of the training with respect to these skills and the ability of participants to absorb the information.

Often the instrument used in this evaluation is a checklist including the relevant skills. Considerable efforts have been made to standardize the items on the checklist. In addition, several training organizations have attempted to standardize the interpretation of each item on the list (e.g., what constitutes satisfactory performance on that item).

However, at the field level there are inconsistencies in terms of the evaluators' criteria for defining competency. Some programs would expect a 100% grade before the trainee would be judged competent on a battery of skills, whereas another organization might consider the person competent if only 50% of the tasks are correctly completed. In some cases, local standards for the delivery of FP services may not exist, in which case international standards can be used.

The assessment of competency is generally more complex than the testing of knowledge (see previous indicator). However, these skills items are often more important to quality of care than the more precisely measurable items of knowledge, and thus they are essential to the evaluation of training efforts.

Indicator

NUMBER/PERCENTAGE OF TRAINED PROVIDERS ASSESSED TO BE COMPETENT AT A SPECIFIC PERIOD POST–TRAINING

Definition

"Trained providers" refers to individuals who have participated in one or more training events.

"Competent" is defined in the previous indicator with regard to a specific skill (e.g., IUD insertion, NORPLANT® implant removal).

Data Requirements

Specification of the skill (e.g., IUD insertion, NORPLANT® implant removal) and established standards for the skill.

Assessment of skills level of trained providers by an expert observer.

Data Source(s)

National standards for service delivery, and checklist and notes of expert observer.

Purpose and Issues

This indicator is useful in determining the retention of skills acquired during training and for identifying possible candidates for retraining. It measures both the adequacy of the training to impart these skills and the ability of the participants to assimilate the information.

If a trained provider does not retain the skills acquired, it is important to explore the reasons, one of which might be a lack of continued practice due to low client load. In fact, this indicator reflects less the quality of the training than the subsequent work environment of the training (e.g.,type and frequency of supervision, demand for the skills).

Conversely, a provider may improve his/her competency by continuously performing the task during the months following the training. In the language of training organizations, the term "proficient" may be used to describe a person who has retained or even improved upon a specific skill as a result of continued practice in the course of routine service delivery.

The assessment of trainees at the given period post–training can be expanded to include a test of their knowledge on topics relevant to family planning. Items on such a test would be similar to those used in assessing "mastery of knowledge" at the end of the course.

Indicator

NUMBER/PERCENTAGE OF TRAINEES WHO APPLY THE SKILLS TO THEIR SUBSEQUENT WORK

Definition

"Skills" must be operationally defined specific to the position and to local standards.

Data Requirements

List of relevant skills, specific to the position of the trained provider; evidence of use of skills.

Data Source(s)

Service statistics regarding numbers of procedures performed; follow–up studies among trained providers.

Purpose and Issues

This indicator is particularly useful in determining the utility of the training to the trainees in their subsequent work. Similarly, it measures the return on the investment made by the organization sponsoring the training. Stated in other terms, this indicator measures the percentage of trained staff that actually work in the field for which they were trained.

If the skill in question is a specific clinical procedure, such as the insertion of the NORPLANT® implant, then service statistics on the number of NORPLANT® implants inserted would serve as evidence that the skill had been applied.

However, in many cases the "skills application" may be less tangible, such as in counseling family planning clients. One would expect improved counseling to result in an increased number of clients and a greater retention rate among clients. However, such concrete measures are also affected by numerous other factors (such as having the trainee transferred out of the FP service to malaria control). In such cases, one might interview the trained service providers to learn what aspects of the training were put into practice on the job.

This indicator reflects less on the quality of the training than on: (1) the adequacy of the selection process (e.g., was the appropriate person sent to the training); and (2) the demand for these skills in a particular service delivery environment. It assumes the system can absorb newly trained providers and give them the service opportunities to use their skills. However, follow–up studies of trained providers may reveal that they are NOT applying their skills to subsequent work if: (1) after the training, they are moved to another position where the skills are inapplicable; or (2) there is low demand for the skill for which they were trained (e.g., vasectomy in certain countries).

COMMODITIES AND LOGISTICS

- Pipeline wastage
- Percentage of storage capacity meeting acceptable standards
- Frequency of stock–outs
- Percentage of service delivery points (SDPs) stocked according to plan
- Percentage of key personnel trained in contraceptive logistics
- Composite indicator for commodities and logistics

The system for providing adequate quantities of contraceptives and related supplies to program service delivery points (SDPs) constitutes a critical element of family planning service delivery operations. Efficient logistics management requires attention to the FIVE RIGHTS: the right quantity of the right quality goods sent to the right place at the right time at the right cost. The role of commodities and logistics system management in transforming program and donor input (in the form of contraceptive commodities and financing) into program output (service access, quality, and program image) is depicted in the conceptual framework sketched in Figure III–1.

In considering indicators for the functional area, attention was focused initially on key areas or subsystems of commodities and logistics system operations. A list of indicators was developed on this basis and presented to the Working Group on Commodities and Logistics in its initial meeting (The EVALUATION Project, 1992a). At this meeting the general consensus was that the development of indicators for commodities and logistics systems was at a relatively early stage in comparison with some of the other functional areas of service delivery. Several CAs represented at the meeting reported that they did not at present have indicators focusing specifically on this functional area.

In reviewing the proposed indicators at the first two meetings of the Working Group and during the intervening period, two problems became apparent. First, a rather large number of factors or characteristics would have to be taken into account in order to assess adequately the status or level of performance of logistics systems. Secondly, a large majority of these factors or characteristics were qualitative in nature, and attempts to quantify them proved difficult and tended to result in key aspects of well–functioning logistics management systems being lost in the effort.

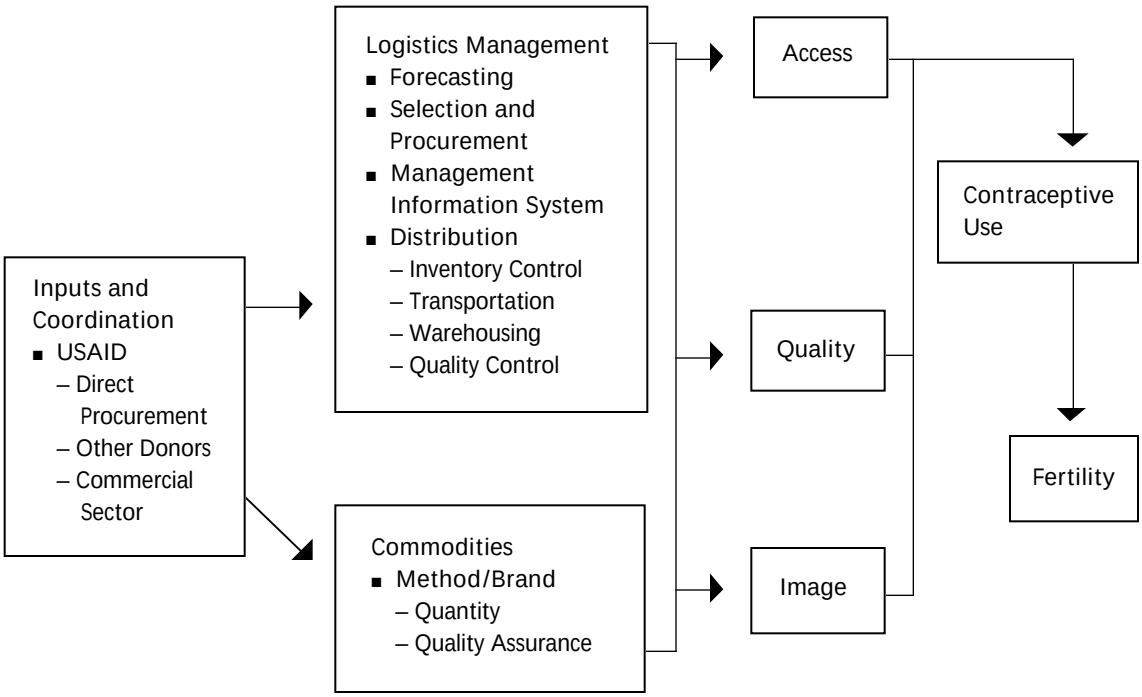
On the basis of these observations, a decision was reached at the second meeting of the Working Group to proceed by: (1) identifying a small number of individual indicators that could be meaningfully quantified; and (2) developing an overall composite index that would summarize in a single measure the many elements or dimensions, both quantitative and qualitative, that were thought to contribute to well–functioning commodities and logistics systems (The EVALUATION Project, 1993a). The model for the development of the composite indicator was the program effort index developed by Lapham and Mauldin (Lapham and Mauldin, 1984). It should be noted that the resulting composite indicator for commodities and logistics is subject to the same concerns regarding level of objectivity as the program effort index.

Working Group discussions emphasized the need for caution in interpreting individual indicators in the commodities and logistics functional area, as the very real possibility exists that indicators might send "conflicting signals;" for example, in order to avoid stock–outs, service providers might be encouraged to ration supplies, thus undermining the ultimate objective of the program. Indicators might also operate in a counter–intuitive fashion as commodities and logistics systems mature. For example, the volume of apparent system wastage might increase as programs improve their ability to track system wastage. Thus, it would appear that individual indicators in this functional area need to be assessed in combination with other indicators and with the composite index in order to be utilized wisely.

In closing, it should be reiterated that work on the development of indicators for this functional area is at a relatively early stage. Future efforts are expected to focus on the feasibility of use of the proposed indicators, issues of measurement objectivity, and the development of indicators for dimensions of commodities and logistics management that may not be addressed sufficiently by the proposed set of indicators.

Figure III–1

Conceptual Framework for the Impact of Commodities and Logistics Activities on Fertility



Indicator

PIPELINE WASTAGE

Definition

The ratio of total contraceptive supplies that are wasted to the amount issued to clients during a specified period of time (e.g., one year).

"Wastage" refers to supplies that expire, are damaged, or are lost.

Data Requirements

Quantities of commodities dispensed to clients during the period, amounts expired or damaged, estimates of losses.

Data Source(s)

Commodities and logistics management information system.

Purposes and Issues

This indicator provides an overall measure of efficiency of the commodities and logistics system, reflecting deficiencies in some of the key components of system operations. The indicator does not, however, provide information on the component or components responsible for observed deficiencies.

A poor logistics system will not have the information required to calculate this indicator.

Some level of wastage is expected even in the best systems. Thus, the indicator is not expected to ever reach zero. However, a small amount of wastage is deemed preferable to the alternative of occasional (or frequent) stock–outs.

Indicator

PERCENT OF STORAGE CAPACITY MEETING ACCEPTABLE STANDARDS

Definition

The percentage of total storage capacity available to the program that meets acceptable standards with respect to temperature, humidity, ventilation, etc.

Data Requirements

Total storage capacity and estimates of capacity that meets standards.

Data Source(s)

Information on total storage capacity should be available from the commodities and logistics management information system.

The percentage of total storage capacity meeting standards will have to be estimated through periodic on–the–spot assessments by senior supervisory staff or other persons with expertise in logistics management.

Purposes and Issues

This indicator provides an overall measure of the adequacy of program storage facilities for contraceptive commodities.

The indicator could be applied at each level of the commodities and logistics system (i.e., at the central, district, and clinic levels) in order to provide a more detailed assessment of the program commodities storage situation at different levels.

It should be noted, however, that storage requirements differ by method; for example, condoms require more storage space per CYP than IUDs. Thus, the indicator gives more weight to condom storage than storage of methods requiring less space.

Evaluation of the adequacy of storage facilities should take into account the following 19–point check list for proper storage:

1. Clean and maintain storeroom regularly.
2. Maintain roof to avoid water leakage.
3. Secure storeroom against water penetration.
4. Ventilate storeroom well.
5. Light storeroom well.
6. Disinfect and spray storeroom for insects regularly.
7. Store supplies away from direct sunlight.
8. Store open cartons of condoms and rubber gloves away from electric motors and fluorescent lights.
9. Stack supplies at least 4 inches (10 cm) off the floor, preferably on pallets made of wood or steel.
10. Stack supplies at least 1 foot (0.3 m) away from any wall and from other stacks of supplies.
11. Separate supplies by lots and in a manner accessible for “First–expiry, First–out” (FEFO) disbursement, counting and general management.
12. Stack supplies no more than 8 feet (2.5 m) high.
13. Arrange cartons so that identification marks and other labels are visible.
14. Issue supplies by carton or box lot.
15. Assure fire safety equipment is available and accessible.
16. Make storeroom accessible at all times.
17. Separate and dispose of damaged and condemned supplies without delay.
18. Store insecticides and chemicals away from other supplies.
19. Store old files, information materials, and related materials separately.

Indicator

FREQUENCY OF STOCK–OUTS

Definition

The percentage of service delivery points (SDPs) that encountered a stock–out of any method/brand during the past 12 months.

Data Requirements

Information on frequency of stock–outs for all SDPs.

Data Source(s)

Commodities and logistics management information system; supervisory and/or staff reports.

Purposes and Issues

This indicator provides a measure of the extent to which SDPs have been unable to serve clients with the full range of authorized contraceptive methods or services during the past year due to inadequate supplies. A more sensitive indicator would be desirable, but alternative specifications are viewed as posing problems in collecting and aggregating information by method and brand.

Under the definition of stock–out adopted for this indicator, a stock–out is deemed to occur when an SDP has no supplies of a particular brand, even though there may be supplies of other brands for the same method.

Caution should be used in interpreting this indicator since family planning workers can avoid stock–outs by rationing supplies.

Indicator

Percentage of Service Delivery Points (SDPs) Stocked According to Plan

Definition

The percent of SDPs having stock levels between their calculated minimum and maximum levels at a given point in time.

Data Requirements

Min/max stock levels for each SDP; actual stock levels at a specified point in time.

Data Source(s)

The commodities and logistics management information system should provide the min/max stock levels.

Data on actual stock levels may be obtained through periodic surveys of SDPs on a sample basis.

Purposes and Issues

This indicator provides an overall measure of the efficiency of the forecasting and distribution components of the commodities and logistics system, but does not provide information on the components responsible for observed deficiencies.

The indicator assumes that a min/max system is in place for SDPs. In such a system, minimum and maximum levels are set based upon amounts issued to clients during prior periods. The minimum stock level is the level below which stocks should not fall without an order having been placed. The maximum stock level is the level above which inventories should not rise under normal circumstances.

Indicator

Percentage of Key Personnel Trained in Contraceptive Logistics

Definition

The percentage of key program staff that have been trained in aspects of logistics management relevant to their role/position.

"Key personnel" are defined as those with significant responsibility for the procurement, storage, distribution and/or dispersement of contraceptive commodities.

Data Requirements

Counts of total numbers of "key personnel" and numbers trained.

Data Source(s)

Program records.

Purposes and Issues

This again is a crude indicator of staff development since it does not provide information on the quality of the training nor the extent to which performance has improved as a result of the training. Nevertheless, it is preferred to indicators such as "Number of persons trained per year" or "Number of training sessions held" since the ideal levels of these indicators will vary according to program needs.

Indicator

Composite Indicator for Commodities and Logistics

Definition

An overall composite index of the state of the system for conducting commodities and logistics activities.

Data Requirements

System–level "scores" for each element and item in the index (see below for details on elements and items comprising the index).

Data Source(s)

Assessments undertaken by senior program staff or others with expertise in logistics management.

Purposes and Issues

As described in the introduction to this section of the Handbook, the purpose of this summary indicator is to provide a basis for assessing commodities and logistics activities at the system level in terms of a sizeable number of relevant characteristics of logistics systems that are somewhat qualitative/subjective in nature and therefore difficult to quantify. The index is intended to be completed by persons knowledgeable in logistics system management as a means of assessing the current status of program commodities and logistics systems in as objective a manner as possible given the inherently qualitative nature of many of the characteristics deemed important to well–functioning systems.

The indicator covers all aspects of logistics systems, and has nine major elements:

- LMIS system,
- forecasting,
- procurement,
- warehousing and storage,
- distribution,
- staffing and training,
- policy issues,
- issues related to donors, and
- organization.

The specific items under each element are shown below. Each item will be scored on a scale, such as 0–1 or 0–3, with a score of 0 indicating a complete lack/absence of the activity and the highest score for the item indicating that the activity takes place completely and correctly. Intermediate scores indicate varying degrees of completeness and quality. All of the items with an element will be summed to yield a total score for the element. The scores for all major groups will be summed to produce a total score for the indicator. Relative weights for different elements and items will be developed by the Working Group on Commodities and Logistics over the course of The EVALUATION Project and will be provided in the update of the Handbook.

The proposed Composite Index elements and indicators are:

1. Logistics Management Information System (LMIS)
   - Has the basic elements of an LMIS: beginning inventory balance, supplies received, supplies issued, ending inventory balance and losses.
   - Cross–checks records: issues data with supplies received, issues data with service statistics, logistics data with survey data, field audits, etc.
   - Uses LMIS system in forecasting and ordering supplies.
   - Provides feedback from LMIS to downstream stations of distribution system.
2. Forecasting
   - Prepares periodic forecasts, by method and brand.
   - Forecasts are linked to procurement time schedules (note: procurement time schedules or cycles vary depending upon factors such as the location of suppliers, source of financing, etc.).
   - Forecasts are incorporated into cost analysis and budget planning.

- Forecasts are verified by comparing them with estimates of actual consumption.
- Forecasts are prepared by internal staff, not outside advisors.

3. Procurement
   - Has names and shipment schedules for all donors and major providers.
   - Coordinates activities of all providers.
   - Has discussions with commercial manufacturers or distributors regarding increased coordination and involvement of commercial sector.
   - Procures from multiple sources.

4. Warehousing and Storage
   - Uses FIFO/FEFO or some other accepted inventory management system.
   - Conducts annual physical inventory.
   - Has and uses appropriate written guidelines for proper storage and handling of supplies.
   - Monitors product quality, including tests of condoms, visual inspection and destruction of expired and damaged stock.
   - Monitors system losses and investigates non–routine losses.
   - Storage capacity is large enough for needs.
   - Storage conditions meet acceptable standards.

5. Distribution
   - Has appropriate distribution plan and schedule for stocking SDPs (e.g., min/max, top up, etc.).
   - SDPs are stocked according to plan.
   - Few SDPs experience stock–outs during the previous year.
   - Does not have excessive pipeline wastage.

6. Staffing
   - The position of Logistics Officer–in–Charge is a dedicated position (i.e., shares no other responsibilities) and is equivalent to other functional unit heads.
   - Percent of key personnel who have been trained in contraceptive logistics.
   - Has appropriate training plan which is used.

7. Policy Issues
   - Has discussions with policy–makers regarding program goals and logistics planning.

8. Organization
   - Has written policies and procedures so that activities can continue appropriately when staff turnover occurs.
   - Has appropriate position within the government agency to which it belongs. This includes the existence of a commodities and logistics unit; unit is respected as having an important role; adequate resources are allocated to the unit.
   - The unit has the ability to coordinate logistics issues within the government, including, for example, the Ministry of Health, Ministry of Finance, etc.
   - Supervision of logistics functions takes place routinely.

9. Donors
   - Has good understanding of the role of the different donors and their future intentions.
   - Maintains close coordination of plans for different donors.

INFORMATION–EDUCATION–COMMUNICATION

- Number of communications produced, by type, during a reference period
- Number of communications disseminated, by type, during a reference period
- Percentage of target audience exposed to program messages, based on respondent recall
- Percentage of target audience who correctly comprehend a given message
- Number of contraceptive methods known
- Percent of audience who acquire the skill to complete a certain task as a result of exposure to a specific communication
- Percentage of target audience exposed to a specific message who report liking it
- Number/percentage of target audience who discuss message(s) with others, by type of person
- Percentage of target audience who advocate family planning practice

Of the different functional areas, information–education–communication (I–E–C) is often the most visible to the general public. Literally millions of people can see and hear the fruits of I–E–C efforts in the form of mass media promotion of family planning. The I–E–C Division is often in charge of community education and outreach efforts, which constitute one of the chief means that the program has for communicating with the target population. Moreover, I–E–C has the potential to directly influence use of family planning, by drawing new users to the services, reinforcing continued contraceptive use among those who have adopted contraception, and altering family size norms.

Not surprisingly, there has been a more sustained interest in trying to evaluate the effectiveness of I–E–C than of the other functional areas of family planning. Two primary designs have been used to this end: a pre– and post–test control group (or non–equivalent comparison group) design and a time series design. Examples of such studies include those by Coeytaux et al., (1987) and Piotrow et al., (1990). It should be stressed that these same designs could be used to evaluate the effectiveness of activities in other functional areas, but historically I–E–C has been a prime focus of such research.

For persuasive communication to achieve its objective, it must evoke in the target person a series of information–processing behavioral substeps (McGuire, 1989). According to classical diffusion theory, the individual passes through an "innovation–decision process" consisting of four functions: knowledge, persuasion, decision, and confirmation (Rogers, 1973). Adapting earlier models, Population Communication Services describes the process using a hierarchy of communication effects model, which traces the steps by which a communications program is expected to have its effect; thus, each step serves as an indicator for evaluating I–E–C interventions (Kincaid et al., 1993).

Given this hierarchy of effects, one might question why this chapter does not include an indicator on "the number/percent that act on an I–E–C message," such as to seek services. Whereas I–E–C activities are key to the knowledge and persuasion

stages, factors other than I–E–C also play a role in the actual decision to seek services and initiate contraceptive use. For example, there must be service sites accessible to the potential client that have trained personnel and adequate contraceptive supplies. Indicators for the implementation stage (shown in the Figure III–3) are described in later sections of this Handbook (specifically, in Chapter VII on service utilization and Chapter VIII on contraceptive use) and are not considered specific to I–E–C.

Most of the indicators for the other functional areas described in this chapter rely on program–based data. By contrast, the majority of the I–E–C indicators require population–based data (percentage of the target audience exposed to FP messages, percentage that report liking the message, percentage that correctly comprehend a given message, etc.). Whereas other functional areas use process and output indicators as the primary measurements for monitoring their activities, the evaluation of I–E–C programs focuses instead on effects. (For a more detailed discussion of this point, see Tsui and Gorbach, 1993). To date, this type of information has been available primarily through special surveys designed for this purpose. However, the questionnaire for DHS III is expected to include a number of I–E–C items that will provide national level data on selected variables.

Figure III–3

Hierarchy of Communication Effects

Proportion of the Intended Audience who:

Knowledge Stage	1. Recalls family planning messages (spontaneous and aided). 2. Comprehends family planning message correctly. 3. Has the knowledge and skills required for effective family planning practices.
Persuasion Stage	4. Likes and approves of family planning messages. 5. Discusses message content and family planning with spouse and friends. 6. Considers population an important problem and supports the family planning program (community norm and support of community influentials). 7. Has a positive image of family planning service providers and modern contraceptives.
Decision Stage	8. Intends to seek family planning information and advice from service providers. 9. Intends to use contraceptives to space children or limit family size.
Implementation Stage	10. Seeks family planning information and advice from service providers. 11. Begins to use an appropriate, effective FP method. 12. Continues to use an appropriate, effective FP method.
Confirmation Stage	13. Recognizes the benefits of contraceptive use, spacing, and small family size for themselves, their family, and their community.

Source: Kincaid et al., 1993

Indicator

Number of Communications Produced, by Type, during a Reference Period

Definition

"Communication" refers to one or more messages packaged as a single item on electronic, print, or other tangible medium (e.g., radio spot, poster, brochure, video, etc.).

Data Requirements

List of items produced in the given period of time (e.g., one year).

Data Source(s)

Administrative records.

Purpose and Issues

This listing constitutes an internal inventory for the program or project, generally done by type of communication. This inventory reflects the capacity of the program to generate materials and thus serves the purpose of creating a sense of accountability among I–E–C staff responsible for production. However, it is a highly crude measure in that it does not reflect the frequency or reach of the diffusion of each communication, nor does it measure the quality of the items produced.

Indicator

Number of Communications Disseminated, by Type, During a Reference Period

Definition

"Disseminated" refers to: (a) the external transmission or distribution of the communications produced via electronic, print, or other media; and (b) the interpersonal activities or public relations events implemented.

Data Requirements

List of communication products disseminated and of activities conducted during a given period of time (e.g., one year).

Data Source(s)

Log books of radio and TV stations regarding the number of broadcasts of each spot or program; data from records of the I–E–C Division on number of posters or brochures distributed to service delivery points (SDP); data from program records at SDPs regarding the number of brochures distributed to clients, educational talks given, outreach visits by program staff, etc.

Purpose and Issues

This indicator measures productivity of the I–E–C Division, specifically the quantity and type of communications disseminated (irrespective of whether anyone sees/hears them, understands them, or acts on them). "Getting the message out" is a necessary (though not sufficient) activity to initiate members of the target population into the hierarchy of effects shown in Figure III–3.

Well–planned I–E–C programs generally have a plan for the diffusion of communications that lists the types of communications and the number for each type to be disseminated. This plan serves as a target to be achieved during the reference period. It is particularly useful to interpret the number of communications actually disseminated in relation to the number targeted. Ideally, one would then ask the question covered by the next indicator, "How successful were these messages in reaching the target population?"

Indicator

Percentage of Target Audience Exposed to Program Messages, Based on Respondent Recall

Definition

"Exposure" refers to an individual's recall of seeing or hearing messages disseminated by the family planning program or other source via electronic, print, or interpersonal channels. The message(s) may be either a specific phrase (e.g., the slogan of an ongoing campaign) or any mention of family planning.

Data Requirements

Count of the channels through which the individual has seen or heard either a specific message or any message about family planning.

Data Source(s)

Survey (preferably with a random sample) of the target population.

Purpose and Issues

Recall of specific messages provides a measure of the reach of a given communications campaign. ("Exposure to" and "reach of" a communications program are equivalent concepts.) For example, in a male motivation project in Zimbabwe, it was estimated that 52 percent of the target population were exposed to project messages (Piotrow et al., 1992).

By contrast, exposure to "any message" about family planning provides a crude but useful measure of the extent to which the public has been informed about family planning via the mass media, whether through promotional messages produced by the government or private family planning association, or through news stories about specific methods. For example, from DHS data it is possible to obtain the percentage of the population exposed to any family planning message. Countries with aggressive media programming on family planning tend to score high on this measure. By contrast, populations that have remained closed to family planning, due either to lack of interest or language barrier, tend to score low.

Two types of recall are frequently used: spontaneous and aided (analogous to the questions on knowledge of family planning methods in the DHS). Specifically, the respondent is asked whether he/she has heard other messages not spontaneously mentioned. This type of "recall" is sometimes labelled "recognition." Since there are usually several messages in an I–E–C campaign, the responses regarding specific messages are: (a) in some cases weighted, more heavily for unaided versus aided recall, and then (b) summed to arrive at a continuous variable measuring level of recall (Kincaid, 1992).

Indicator

Percentage of Target Audience who Correctly Comprehend a Given Message

Definition

In operational terms, the percentage of persons who, having heard a specific message, are able to correctly paraphrase the main idea.

Data Requirements

Answers from respondents in either a pre–test or in a post–diffusion survey.

Data Source(s)

Survey (preferably with a random sample) of the target population.

Purpose and Issues

This indicator is useful in ensuring that the messages being disseminated are indeed comprehended by the target population. Ideally, all messages should be tested for comprehension (as well as other dimensions) prior to final production. However, even if they pass a pre–test based on a small, non–representative sample of the population, it is useful to assess comprehension once the messages are actually in circulation among the target audience.

It is important to collect this information by interviewing a series of individuals in private (the usual format for a survey). By contrast, focus groups are not a useful means of obtaining this information. If, for example, only one person in the group knows the correct response and he/she gives it, this immediately contaminates the rest of the data collection procedure.

Indicator

Number of Contraceptive Methods Known

Definition

The number of contraceptive methods that a survey respondent is able to spontaneously mention or is able to recognize when mentioned by the interviewer.

Data Requirements

On the DHS and similar surveys, the respondent is asked to name all the contraceptive methods he/she has heard of. For each method not mentioned, the interviewer names the method and asks if he/she has heard of it. It is possible to differentiate the number given from *spontaneous mention* versus *prompted recall*. More often, the two are combined and reported as *the number of methods known*.

Data Source(s)

Survey of the target population (e.g., DHS).

Purpose and Issues

This indicator reflects the extent to which the target population is aware that methods exist to prevent pregnancy. It is reported alternatively as the percentage who know at least one contraceptive method or the mean number of methods known. If a country is interested primarily in the use of modern methods, the indicator can be limited to this category of methods.

In countries with moderate to high levels of contraceptive prevalence, close to 100% of the adult population have heard of at least one method, a reason that this indicator is of little interest to many population researchers. By contrast, in countries with emerging family planning programs, knowledge of contraceptive methods is often the first population–based indicator to "respond" to FP program interventions (i.e., show change at the population level). Whereas increases in actual contraceptive use may take years to achieve, changes in knowledge can occur within months of the initiation of an intensive I–E–C effort.

One major criticism of this indicator is that it is labelled "knowledge," but it does not measure whether the respondent has a meaningful understanding of the method(s) in question. The respondent can be credited with "knowing" a method by nodding appropriately as the interviewer reads down the list of a dozen modern and traditional methods, even if he/she has no idea what the contraceptive looks like or how it is used. A more appropriate label for this indicator would be awareness of contraceptive methods, but researchers worldwide have come to refer to this block of questions on DHS surveys as knowledge of contraceptive methods.

One might question the validity of this indicator on the grounds that a respondent might claim to "recognize" more methods than he/she had actually heard of to avoid looking ignorant to the interviewer. Although this situation may occur, there are two reasons to believe that this indicator does in fact provide a valid reflection of levels of awareness regarding contraception. First, within a given country (especially those with low to moderate levels of prevalence) respondents tend to know the methods most commonly in use, but deny knowing less common ones (e.g., spermicides, diaphragm, or vasectomy, depending on the country). Second, the percentage knowing at least one method tends to increase monotonically over time, a finding one would expect as populations become more familiar with the concept of family planning.

Indicator

PERCENT OF AUDIENCE WHO ACQUIRE SKILL TO COMPLETE A CERTAIN T RESULT OF EXPOSURE TO A SPECIFIC COMMUNICATION

Definition

"Acquiring" the skill to do a task implies that the individual was not able to do it correctly prior to seeing or hearing the communication. "Skill" refers to behavior needed to correctly use a specific contraceptive (e.g., taking the pill in the correct sequence, putting on a condom, checking the strings of an IUD, etc.).

Data Requirements

Definition of criteria for the correct performance of the task; verbal description of how to perform task or actual demonstration, before and after exposure to the communication.

Data Source(s)

Interviews with members of target population exposed to the messages and/or actual observation of the skill performed.

Purpose and Issues

This indicator is intended to measure the effectiveness of a given communication in teaching a skill, assuming that is the purpose of the communication.

Ideally, the researcher will be able to observe members of the intended audience actually performing the task. Verbal reports are less reliable than actual observation. On the one hand, respondents might claim greater facility at doing a task than they actually have; on the other, less articulate respondents might be better at actually doing a task than explaining how it might be done.

However, in some cases, the choice of approach may be dictated by the circumstances (e.g., if the skill is to learn how to use a condom correctly; the "compromise" measure might be to demonstrate correct usage on a plastic model).

While this indicator is important as part of the conceptual hierarchy of effects, it is among the most difficult to apply in practice. Even if the respondent is able to "demonstrate" the skill in a simulated exercise, he/she may not apply it in everyday living. In addition, the field team needed to assess whether an individual has acquired a given skill may require a higher level of training and/or clinical expertise than the typical interviewer would have. Collecting data on this indicator represents a special challenge, and thus it is not widely used at the field level.

Indicator

Percentage of Target Audience Exposed to a Specific Message who Report Liking It

Definition

"Liking" is defined as a positive emotional response to a given communication, based on self–report of respondent.

Data Requirements

Responses to questions asked on surveys, in–depth interviews, focus groups, etc.

Data Source(s)

Survey (preferably with a random sample) of the target population; focus groups.

Purpose and Issues

This indicator gives a sense of the affective reaction of respondents to the communications produced by the I–E–C program. Communications that create a positive response among members of the target population are more likely to shape positive attitudes toward the proposed behavior (e.g., contraceptive use).

It should be noted, however, that there may be a courtesy bias in response to this question, especially if respondents believe the interviewer works for the family planning program.

The DHS survey includes a question on approval of family planning messages, which generally refers not to a specific message, but rather to the idea of using radio/TV as a channel for disseminating such messages. Thus, this piece of information is available for numerous countries around the world. Although different from "attitude toward family planning," the two tend to be highly correlated.

Indicator

Number/Percentage of Target Audience Who Discuss Message(s) With Others, by Type of Person

Definition

"Discussing the message" refers to any conversation subsequent to exposure to the communication in which the communication (spot, brochure, song, etc.), its characters, or messages are mentioned. "Type of person" includes spouse, partner, relative, friend, etc.

Data Requirements

Number of persons who discussed the family planning messages with others, as a percentage of:

- those who heard/saw the messages in questions, or
- those interviewed.

Data Source(s)

Survey (preferably with a random sample) of the target population.

Purpose and Issues

This indicator measures the extent to which one or more messages from the mass media generate further interpersonal communication.

Within the communication field, it is often stated that mass media are useful to create awareness and increase knowledge, but that interpersonal communication plays a vital role in bringing about actual behavioral change. Whereas it has also been shown that media can have a direct effect on behavior, campaigns that generate substantial interpersonal communication may result in an even greater level of behavioral change (first, because of the social support that may be generated for the idea; second, because the message may be transmitted to others who did not hear it from the original source).

The "spin–off effect" is not necessarily positive. A campaign judged to be in poor taste might create great controversy, much of which could be negative. (On the other hand, some would argue that publicity—any publicity—is useful.)

Indicator

Percentage of Target Audience who Advocate Family Planning Practice

Definition

In operational terms, the percentage of persons who either recommend family planning practices to their friends and relatives, including possibly taking them for family planning services, and those who speak out or provide some public testimonial in support of a program, such as participating actively in community events, encouraging support from community, political, or health leaders for enhanced services.

Data Requirements

Number of persons that recommended the use of family planning to relatives and friends or the number of persons who participated in public events or spoke up publicly in support of family planning progress as a percentage of:

- those who heard/saw the messages in question; or
- those interviewed.

Data Source(s)

Survey (preferably with a random sample of the target population) or surveys of specific groups or organizations which have worked in the programs, such as nurse/midwives associations, medical associations, women's organizations, or other organized social or professional institutions with members in the community.

Purpose and Issues

This indicator measures the extent to which support for family planning is no longer a private, individual, even taboo practice among individuals, but rather, is perceived as a community norm and a valuable part of community activities.

Within the field of communication, the willingness of an adopter or supporter of a program to bear public witness of his or her support and commitment is a measure of the depth of personal commitment. Persons who have spoken out publicly in support of a measure (whether family planning, smoking cessation, or use of drugs) are less likely to discontinue their new practice. At the same time, the act of speaking out publicly reinforces personal support for a particular practice. Also, expressions of support from satisfied users will enhance practice within the community, and expressions of public support within community forums will increase national and community support for programs in the long run. This type of behavior has not been measured previously, but can be an increasingly important indicator in the future to measure local financial and political support for family planning and related programs.

RESEARCH AND EVALUATION

- Presence of an active research and evaluation unit
- Extent of use of a service statistics system
- Conduct of periodic household and/or special purpose surveys and studies
- Conduct of operations research (OR)
- Regular conduct of process evaluations
- Conduct of effectiveness, efficiency, and impact evaluations
- Use of research and evaluation results for program modification
- Dissemination of research and evaluation results

Previous research has demonstrated a strong association between the conduct of program–related research and evaluation and family planning program performance (Lapham and Mauldin, 1984; Entwisle, 1989; Mauldin and Ross, 1991). In their well–known program effort index, Lapham and Mauldin (1984) include "record keeping and evaluation" as one of four major components or categories of indicators in the overall program effort score. The relevance of such activities for program performance, and in particular the close relationship between the conduct of evaluation activities and indicators of program management effectiveness, has been empirically verified by Entwisle (1989).

Research and evaluation is viewed in The EVALUATION Project conceptual framework as an essential supporting activity to effective management of family planning programs. Although research and evaluation activities are not directly involved in service delivery, program capacity to generate and effectively utilize research and evaluation findings is thought to contribute in important ways to effective management of service delivery. The presumed supporting role of research and evaluation to service delivery, and in particular to effective program management, is depicted in the conceptual framework presented in chapter I of this Handbook.

Research and evaluation undertakings contribute information as input to management decision–making through such activities as:

- assessment of progress in program implementation;
- the identification of factors contributing to observed program deficiencies and testing of alternative corrective actions;
- the determination of the degree of impact that the program has had in the target population in terms of influencing relevant outcome; and
- assessment of the level of program effectiveness and efficiency.

The indicators presented for the research and evaluation functional area are intended to provide a basis for assessing program capacity to undertake research and evaluation and utilize the products of such efforts to improve program functioning and performance. The indicators fall along two main dimensions:

- program capacity to generate data/information relevant to management information

needs using appropriate data collection systems, study designs, and methodologies; and

- the degree of meaningful utilization of research and evaluation output in program planning and management decision–making.

The specific set of indicators chosen reflects the view of The EVALUATION Project as to what would comprise a comprehensive monitoring and evaluation program; that is, the issues to be addressed and the data collection systems and types of data necessary to fully inform management decision–making with respect to issues of input, process, output, outcome/impact, effectiveness and efficiency. It is recognized, however, that the priority management issues faced by family planning programs will vary depending upon the type of program (public sector, NGO, CBD) and the stage of program development. Accordingly, needs for and capacity to undertake research and evaluation will vary from program to program and will tend to evolve over the course of program development.

For example, programs in the start–up phase would ordinarily be focused more on monitoring inputs, processes and outputs than on issues of impact and efficiency. Similarly, more mature programs, having achieved some degree of success in meeting program output and intermediate outcome targets, would normally be expected to devote more attention to issues of effectiveness, efficiency, and impact. Thus, it is not intended that all of the indicators presented will necessarily be used in a particular setting, except perhaps in the case of relatively mature programs. Accordingly, the selection of indicators for use in a particular setting should be guided by what might be reasonably expected given a program's stage of development and priority management information needs.

As a result of the attempt to identify indicators for the functional area based upon The EVALUATION Project's perception as to what constitutes a comprehensive monitoring and evaluation system, the setof indicators proposed might be viewed as too heavily weighted toward the conduct of different types of evaluation undertakings. This issue will be revisited in the update of this Handbook based upon field experience and feedback during the next few years. Possible remedies might be to either collapse indicators pertaining to the conduct of different types of evaluation undertakings into a single index, or to weight the indicators in a manner such that each indicator has appropriate influence in an overall "score" for the functional area.

Indicator

Presence of an Active Research and Evaluation Unit

Definition

An appropriately staffed unit with principal responsibility for program–related research and evaluation that undertakes relevant activities on a regular basis has been established within the program.

"Appropriate staff" are defined as persons with training and/or experience in relevant fields or disciplines (e.g., demography, operations research, social science research methods, statistics, etc.).

Data Requirements

Evidence that a research and evaluation unit has been formally established and is functioning.

The presence of a research and evaluation unit on the program organizational chart and staffing patterns indicating the formal assignment of appropriate personnel to the unit constitute evidence of the existence of a unit.

The following types of undertakings may be considered evidence of regular activity: (1) periodic estimation and analysis of the level of contraceptive use or prevalence from service statistics and surveys; (2) implementation of or participation in surveys; operations research and/or other special studies; (3) preparation of technically sound reports on program accomplishments; (4) preparation of reports summarizing (for program management) relevant national and sub–national data collected through various sources (e.g., population censuses, vital statistics, surveys); and (5) review and synthesis of results of relevant technical reports and studies produced by other (non–program) researchers for program management.

Data Source(s)

Program records and documents; research and evaluation reports.

Purpose and Issues

The present indicator provides a measure of whether a formal research and evaluation unit has been established and whether relevant types of activities are undertaken on a regular basis. Beyond establishing a link between research and evaluation activity on the one hand and program performance on the other, prior research suggests that the location of a research and evaluation capability within family planning programs plays an important role in improving program planning and management performance (Lapham and Mauldin, 1984; Entwisle, 1989; Mauldin and Ross, 1991). The benefits of an internal research and evaluation capability, as opposed to reliance on institutional resources and capabilities external to the program, presumably lie in three areas: (a) an enhanced ability of program management to control the research agenda, thus focusing attention and resources on key issues and questions; b) better lines of communications between the researchers/evaluators and program managers, thus increasing the likelihood that research and evaluation results will be disseminated to key decision–makers; and (c) "ownership" of the research and evaluation efforts and findings, thus diminishing the likelihood that research and evaluation activities will be viewed as threatening and increasing the likelihood that the findings will be used to improve program performance.

Balanced against these are two considerations. First, in many developing country settings, the technical skill levels of external institutions (e.g., university researchers and private sector research firms) may be higher than those of program research and evaluation personnel. Secondly, there is a question of whether vested interests in making the program "look good" will prevent an objective evaluation from being conducted and sensitive issues from being addressed by internal personnel.

In their scoring system for the program effort index, Mauldin and Ross (1991) advocate assigning only partial "points" if only an external research and evaluation capability exists in a particular setting. In view of the perceived benefits of an internal capability outlined above, we are inclined to concur with this approach, although perhaps additional credit might be given if external research and evaluation capabilities are effectively used by the program.

With regard to the question of objectivity, it is felt that the lack of objectivity in undertaking and utilizing the results of research and evaluation efforts will be adequately reflected in other indicators under the Management and Research and Evaluation functional areas.

Indicator

EXTENT OF USE OF A SERVICE STATISTICS SYSTEM

Definition

A service statistics system is routinely used by program staff for program management decision–making, supervision, and program evaluation purposes.

Data Requirements

Evidence that the service statistics system is used in undertaking program–related research and evaluation. The following are illustrative of the types of activities that might be considered evidence of utilization: (1) use in staff and facility performance appraisal; (2) use in setting service targets and assessing program performance against targets; and (3) use in decisions regarding proposed changes in program strategies or operations.

Data Source(s)

Program records and documents; interviews with program managers and other intended users of service statistics to determine the extent of meaningful use; evaluation reports in which service statistics are used.

Purpose and Issues

As noted in the discussion of indicators for the management functional area, the establishment of a service statistics system is viewed as essential to effective program management. The intent of the present indicator is to go beyond the mere establishment of a service system to attempt to capture the extent of meaningful use of service statistics by program staff in carrying out management–related duties and responsibilities, including program planning and evaluation. In a number of countries and/or programs, service statistics systems have been implemented, but the data are of questionable quality and are not meaningfully utilized. The present indicator attempts to measure the qualitative dimension of the level or degree of meaningful utilization of the service statistics system.

One of the more useful aspects of service statistics systems for program evaluation purposes is the opportunity to give feedback to service delivery points that they provide; for example, by making note of accomplishments (or lack thereof) against annual service targets (e.g., number of family planning clients seen, number of new acceptors, etc.). Here, however, caution must be exercised to ensure that service statistics and service targets are not used in an overly mechanical manner, for there is a real danger that too strong an emphasis on meeting targets will result in program staff focusing on meeting their "quotas" to the detriment of service quality.

The assessment of effective program utilization of service statistics should also take into account the purposes/issues for which service statistics are most appropriate. Service statistics provide information on services provided at program facilities and on clients who have come into contact with such facilities, and are thus useful for addressing questions and issues related to program processes and outputs. Service statistics are less appropriate, however, for measuring population–based outcomes, especially in settings where multiple sectors (public, private, and commercial) are heavily involved in the provision of family planning services. Population–based surveys and other population–based studies are more appropriate sources of data for measuring population–based outcomes. Thus, programs that use service statistics data inappropriately should be evaluated less favorably on this indicator than programs whose use of such data reflects an awareness of the strengths and limitations of service statistics.

Indicator

Conduct of Periodic Household and/or Special Purpose Surveys and Studies

Definition

Program–relevant periodic household and/or special purpose surveys and studies have been undertaken within a given reference period (e.g., the last 3–5 years).

Data Requirements

Evidence, in the form of survey or study outputs, that periodic household and/or special purpose surveys have been conducted within the specified reference period.

Data Source(s)

Program records and documents; survey or study reports or other outputs.

Purpose and Issues

Because of the inherent limitations of service statistics, effective research and evaluation support to program management requires that supplementary information be available on a periodic basis. Two particular types of essential information for which service statistics are not an appropriate source are: (1) population–based data on behaviors, knowledge, attitudes, and practices relevant to the use of contraception in the target population for program services as a whole (that is, among non–clients as well as clients); and (2) detailed studies of behaviors, perceptions, and needs of program clients as they relate to program services.

Sample surveys and other special studies based upon data sources other than program records and/or service statistics provide a valid and cost–effective means of generating such information. In addition to large–scale household surveys, relevant data collection strategies include small–scale KAP surveys, exit interviews with clients, and focus group discussions. The present indicator is intended to measure the extent to which the program is able to generate (or obtain access to) such data as needed.

In assessing programs on this indicator, it should be recognized that family planning programs sometimes do not possess sufficient personnel or logistic resources to undertake certain types of data collection activities on their own; for example, the conduction of large national sample surveys such as the Demographic and Health Surveys (DHS). More often than not, data collection for such large–scale data collection efforts is undertaken by the national statistics office or similar specialized agency in collaboration with agencies for which the survey results have programmatic implications (e.g., the national family planning program or ministry of health). In such instances, credit should be given to programs for meaningful participation in the activity; for example, collaboration in the specification of survey or study content or measurement objectives, design of the questionnaire and tabulation plan, analyses of data, and report preparation.

Similarly, programs might also receive "credit" on this indicator for survey or study efforts for which relevant responsibilities have been "contracted out" to researchers or institutions external to the program. In some instances, contracting data collection and/or analysis activities to external organizations provides a cost–effective means of supplementing in–house research and evaluation capabilities, and in some cases the only available means of obtaining needed information. However, in order to ensure that the survey or study undertaking is responsive to program information needs, meaningful program participation in the study effort is highly desirable. Meaningful program participation might be defined in terms of the following types of involvement: specification of survey or study measurement objectives and statistical precision requirements, development of the study design, participation in the design of the questionnaire and dummy tables, supervision of the contractor during fieldwork, and/or participation in analysis and report preparation.

Indicator

CONDUCT OF OPERATIONS RESEARCH (OR)

Definition

One or more operations research (OR) studies have been conducted by the program within a given reference period.

"Operations research" studies refer to applied studies examining the causes of and possible solutions to observed program operational problems (Blumenfeld, 1985; Fisher et al., 1991).

Data Requirements

Evidence, in the form of reports or other outputs of OR studies and/or evidence of staff involvement in ongoing studies that operations research studies have been conducted within the specified reference period.

Data Source(s)

Program records and documents; OR reports or other research output.

Purpose and Issues

Operations research provides a means of systematically examining the factors underlying observed program operational deficiencies and the relative merits of alternative corrective actions. Program capacity to conduct and make use of OR studies in modifying program strategies and operations to improve effectiveness or efficiency is viewed as an essential element of effective research and evaluation support of program management. This indicator is intended to measure program activity in undertaking OR studies.

Because the OR process itself is viewed by many practitioners as being as important to improving program operational efficiency as the research results, it is viewed as essential that programs themselves take the lead in conducting OR studies, if not being fully self–sufficient in the OR process and in relevant research methods and approaches. Thus, it is proposed that programs receive the "full score" on this indicator only where the program has demonstrated institutionalization of OR through the conduct of one or (preferably) more OR studies in which the program has been the lead player in the study effort; that is, with minimal levels of external assistance.

Like several other indicators in this section of the Handbook, the present indicator provides only a summary measure of program capacity to conduct OR studies. The indicator does not take into account other relevant dimensions of program OR capacity such as the technical quality of the research undertaken, the degree to which the information generated fills existing information gaps, or the extent to which the findings of OR studies are meaningfully utilized by the program. More detailed assessments of program OR capabilities would require that additional indicators be measured.

For reference purposes, The EVALUATION Project Working Group on Operations Research in its initial meeting proposed the following indicators as relevant for more in–depth evaluations of OR capabilities at the program level (The EVALUATION Project, 1992d):

- development of an OR agenda;
- degree of usefulness of the information generated;
- degree of institutionalization of OR as a management tool (a dimension which is at least partially incorporated into the specification of the present indicator);
- the level of program change resulting from OR efforts (a similar indicator, pertaining to the utilization of research and evaluation results generally, as opposed to OR results specifically, is presented later in this section of the Handbook);
- the quality of the actual research; and
- the level of effort expended to utilize findings from OR studies to improve service delivery.

A full listing of the indicators drafted at the first meeting of the Operations Research Working Group is included in Appendix C.

Indicator

Regular Conduct of Process Evaluations

Definition

Process evaluations have been undertaken by the program on a regular basis within a given reference period.

"Process evaluations" refer to undertakings that assess the extent to which scheduled activities occur on time, in the manner expected (that is, according to specified program standards), and at the expected cost, as well as the extent to which expected program outputs have been realized.

Data Requirements

Evidence, in the form of evaluation reports or other evaluation effort outputs, that process evaluation efforts have taken place on a regular basis within the specified reference period.

Data Source(s)

Program records and documents; evaluation reports and outputs.

Purpose and Issues

Process evaluation is concerned with assessing the extent to which the program is on track or is moving forward in the expected manner. In this sense, process evaluation is a program monitoring strategy, and the term "monitoring" is used by many programs in this context. The sense of the present indicator, however, is that process evaluation entails the comparison of results and accomplishments against plans or expectations, and thus goes beyond merely keeping track of or counting program outputs. The indicator is intended to measure the extent to which a program monitors progress against plans or expectations on a regular basis.

The definition of "regular basis" for the indicator may vary somewhat from program to program, but should conform to the planning cycle used by the program. For example, if annual workplans are used, process evaluation would normally be undertaken on an annual basis in order to both assess performance during the past year or planning cycle and provide input into developing the work plan for the next year or cycle. The systematic and regular conduct of process evaluations is the key dimension of the proposed indicator, and accordingly "less than regular" process evaluation schedules or evaluation schedules that do not complement the program planning cycle should be reflected in lower ratings on this indicator.

Indicator

CONDUCT OF EFFECTIVENESS, EFFICIENCY, AND IMPACT EVALUATIONS

Definition

One or more effectiveness, efficiency, and impact evaluations have been undertaken by the program within a given reference period.

"Effectiveness evaluation" refers to evaluation undertakings designed to measure the extent to which the program has produced expected intermediate–term outcomes at the population level.

"Efficiency evaluation" refers to efforts to assess the relationship between program inputs and outputs.

"Impact evaluation" refers to evaluation activities that seek to measure long–term changes in fertility or other intended population–level outcomes that are attributable to the program.

Data Requirements

Evidence, in the form of evaluation reports or other outputs of evaluation efforts that one or more of these types of evaluation have been conducted by the program within the specified reference period.

Data Source(s)

Program reports and documents; evaluation reports or other outputs.

Purpose and Issues

While process evaluations provide information for program management purposes on a routine basis, effective management requires periodic information on issues that cannot be meaningfully addressed by the types of data that are normally used in process evaluations. Among these are issues relating to program effectiveness, efficiency, and impact.

Effectiveness evaluations entail the assessment of the extent to which the program is achieving satisfactory progress toward the achievement of expected population–level outcomes. It is the focus on population–level intermediate outcomes or effects that distinguishes effectiveness evaluations from process evaluations, which involve the assessment of progress toward achieving intended program–level outputs. For example, the achievement of program targets for numbers of new acceptors would be the appropriate focus of process evaluation efforts, while the achievement of increased contraceptive prevalence in the target population would be the proper focus of effectiveness evaluations.

The two types of evaluation are connected, however, in that achievement of intended program–level outputs is expected to result in, or at least contribute to, the achievement of population–level outcomes.

Efficiency evaluations refer to assessments of the relationship between program inputs and outputs. The focus of such evaluation is the measurement of the per–unit cost of program output. The cost per CYP or per new acceptor is an example.

It should be noted that while program efficiency has been conventionally measured as defined above, it is also useful to examine the issue of efficiency in terms of population–level outcomes. The rationale for this is twofold. First, population–based outcomes constitute the ultimate objective of family planning programs. Secondly, measures of efficiency based on program outputs as opposed to population–level outcome may not point to the same conclusions. For example, distribution of oral contraceptives through a CBD program may be accomplished in a highly cost–efficient manner, but unless oral contraceptives are widely understood and accepted, little change may result in contraceptive prevalence or fertility rates. Thus, measures of program efficiency based upon both outputs and outcomes provide useful information for management decision–making.

Impact evaluation refers to assessments of the ultimate outcomes or net result(s) of the program, both positive and negative, at the population level. Impact evaluation efforts are aimed at providing answers to the question, "How has the program changed relevant ultimate population outcome indicators (e.g., fertility levels, contraceptive prevalence, satisfaction of reproductive preferences, infant/child death rates, etc.)?"

Answers to the types of questions addressed by effectiveness, efficiency, and impact evaluations

cannot normally be derived from routine sources of data (e.g., service statistics and MIS data), and thus special efforts are usually required to compile and/or analyze data in connection with the evaluation effort. In addition, the measurement of program impact requires attention to the issue of attribution to the program (see the introduction to Chapter IX for further discussion of this issue).

Since these types of evaluations are intended to measure medium– to long–term outcomes, they do not need to be conducted as frequently as process evaluation. Accordingly, some flexibility should be allowed in defining the reference period for the indicator, taking into account the age or level of maturity of the program, the recency of major changes in program policy and/or modes of operation, and other factors that influence the need for the types of information produced by such evaluation efforts.

It should be noted that the indicator is intended as a crude summary measure of the level of program activity in the conduct of non–routine program evaluation. As such, the indicator does not address issues such as the technical quality of evaluation efforts nor the extent to which evaluation results are meaningfully utilized by program management. Assessment of these dimensions would require the inclusion of additional indicators.

Indicator

Use of Research and Evaluation Results for Program Modification

Definition

Research and evaluation findings have played a significant role in decisions regarding program policies, strategies, operational procedures, etc., within a given reference period.

Data Requirements

Evidence, either in written or verbal form, of instances in which research and evaluation findings played a key role in program decisions during the specified reference period.

Data Source(s)

Program records and documents; interviews with program managers and key decision–makers.

Purpose and Issues

This indicator is intended as a measure of the extent to which findings from research and evaluation efforts have been meaningfully utilized by program management in arriving at decisions regarding key program elements. Although the focus of the indicator is on the use of research and evaluation findings to motivate or justify changes in program strategy or operations, the utilization of research and evaluation results to support decisions to maintain existing strategies or operational modes are equally relevant.

In terms of measurement, the key is the establishment of a linkage between decisions to make (or not to make) program changes on the one hand and results of research and evaluation undertakings on the other. In some instances, the linkage will be clearly established by written documents detailing changes in program strategy and/or modes of operation in which research and evaluation results are cited as at least a contributing factor in decisions to make changes. More often, linkages between research and evaluation findings and program decisions will have to be established through assessments of the extent to which program strategies and operations are consistent with study findings and interviews with key management personnel in which the degree of utilization of research and evaluation findings is assessed.

Indicator

DISSEMINATION OF RESEARCH AND EVALUATION RESULTS

Definition

Research and evaluation results are regularly disseminated to key external audiences (e.g., key government agencies, the news media, the research community, etc.).

"Dissemination" refers to the formal communication of program–related research and evaluation findings through such channels as research briefs, publications, workshops, conferences, news releases, etc.

Data Requirements

Evidence of events or instances in which research and evaluation findings have been disseminated.

Data Source(s)

Program records and documents; evaluation reports and other outputs.

Purpose and Issues

Dissemination of research and evaluation results serves a number of essential functions in family planning programs. Dissemination or communication of findings to key audiences internal to the program, normally mid– to upper–level program managers, constitutes the principal means of connecting the research and operations units of programs, which are often separated administratively and tend to have less than optimal lines of communication between them. Dissemination to external audiences (for example government agencies, legislative bodies, the news media, the research community, and external donors) provides a useful means of publicizing program accomplishments, heightening the visibility of the program, and calling attention to issues requiring action.

Although both internal and external dissemination are important, the focus of the present indicator is on dissemination to external audiences. It is assumed that lack of internal communication between research and evaluation units and program management will be reflected in other indicators presented for the Management functional area (see Section A of this chapter). Thus, the present indicator is intended to provide a measure of the extent to which programs routinely disseminate relevant research and evaluation findings to external audiences.

For measurement purposes, it is recognized that program resources for dissemination are often limited. It is also the case that not all research findings are worthy of wide dissemination. At a minimum, programs should have a policy to routinely review the results of research and evaluation efforts to assess their "newsworthiness," as well as perhaps an annual or semi–annual summary report describing research and evaluation efforts undertaken and providing a brief synopsis of results.

Chapter IV

Family Planning Service Outputs

A. Accessibility

B. Quality of Care

C. Program Image

Chapter IV

Indicators to Measure Family Planning Service Outputs

The service delivery operations outlined in the previous chapter are expected to have a direct effect on the three dimensions of service delivery, described in this chapter as the service outputs: access, quality of care, and program image. The indicators of service outputs presented in this chapter are based largely on client perspective and answer to the question: what does the potential family planning client find when he/she seeks contraceptive services?

In the past, program evaluation has tended to focus on measures of service utilization and contraceptive prevalence. All too often, the program itself is treated as a "black box." The results of such evaluations show the extent to which the program achieves expected results, but give little indication of why it has succeeded or failed. Assuming the purpose of evaluation is to improve service delivery in the future, lack of attention to the inner workings of the program is a serious shortcoming.

In fact, these topics are coming under increasing attention from both researchers and practitioners. Advances in survey methodologies and geographic information system technology are opening new horizons for assessing the adequacy of physical access to family planning services. At the same time, quality of care has become a central focus of service providers in efforts to improve the delivery of family planning services. As a result, researchers are scrambling to find methods of quantifying what previously was thought to be too subjective for systematic assessment. Of the three categories of service outputs, program image has to date received the least amount of attention, but nevertheless remains a key determinant of utilization.

As a preface to this chapter, it would be useful to address several conceptual issues, the resolution of which underlies the development of indicators in this area. One such issue is the interrelationship between access and quality in the delivery of family planning services. In working group discussions on this topic, the argument has repeatedly surfaced that a client cannot be satisfied with services unless they are accessible. Thus, it is argued, access to services must be included in the definition of service quality.

However, many examples may be cited of family planning programs that have succeeded in providing reasonable access to services to a large proportion of the program's target population, but the services are not well received and are under–utilized.

In order to clarify the relationship between these two concepts, it is useful to distinguish factors that determine whether an individual interested in obtaining family planning services actually "gets to the door" (that is, overcomes the different types of barriers that family planning clients often face when seeking services) versus what he or she finds once "inside the door" (e.g., the cleanliness of facilities, the administrative efficiency and interpersonal communications skills of service providers, reasonable waiting times, etc.). Note that "door" is used here in a figurative sense; it is not intended to refer exclusively to clinic–based services, but may also refer to the "door" of a CBD worker's house or of the local pharmacy. Both access and quality are thought to be important determinants of client satisfaction (and presumably contraceptive acceptance and continuation). However, different program management responses would be required to address problems encountered in "getting clients to the clinic" as compared to "keeping them wanting to come back." Thus, it is analytically useful to view access and quality as distinct (but complementary) elements of the family planning supply environment.

Another dimension of service output is program image. A family planning program's public image refers to the degree to which public percep-

tions about the program are favorable. A number of factors go into the creation of this image, including general population perceptions about family planning and public and private sector service providers, as well as perceptions (both founded and unfounded) about the family planning services themselves. Since the latter are influenced by the other dimensions of what are referred to in this chapter as service outputs (i.e., access to and quality of services), the program image is also viewed as a service output to the extent that it is (at least partially) controllable by program management. For example, many programs attempt to improve their program image by improving the quality of services provided. Programs may also attempt to influence public perceptions as to the legitimacy of family planning programs through I–E–C efforts (as discussed in Chapter III).

Although program image, access, and quality are often correlated (e.g., programs that offer high quality services and that are readily accessible often have a favorable program image), it is sometimes the case that improvements in a program's image lag behind improvements on the other service outputs due to lingering misperceptions about the program. Accordingly, program image is conceptualized as being an important output of family planning program efforts that warrants the attention of program management over and above attention given to issues of access and quality.

ACCESSIBILITY

Illustrative Indicators

- Number of SDPs located within a fixed distance or travel time of a given community (i.e., service density)
- Cost of one month's supply of contraceptives as a percentage of monthly wages
- Restrictive program policies on contraceptive choice
- Percentage of the population who know of at least one source of contraceptive services and/or supplies
- Percentage of non–use related to psycho–social barriers

The conceptualization and measurement of access to family planning services and the relationship between service accessibility and family planning outcomes (e.g., contraceptive acceptance and continuation) have received considerable attention over the years by family planning program managers and researchers. Despite the widely acknowledged importance of accessibility as a key feature of the supply environment, however, there is little consensus as to the most appropriate way(s) to measure the concept.

Much of the previous research in this area has focused on one aspect or dimension of accessibility: geographic (or physical) accessibility. In this context, accessibility refers to the degree of difficulty in reaching or obtaining family planning services. A variety of measures pertaining to the distance to supply and service points, the time required to reach these points, and the density of service/supply points within a specified geographic area have been proposed in the literature (Chayovan et al., 1984; Hermalin and Entwisle, 1985 and 1988; Hermalin, et al., 1992; Tsui and Ochoa, 1992). While the evidence to date tends to confirm the relevance of geographic proximity to family planning services as an important determinant of contraceptive use, the strength of the relationship between proximity and contraceptive use in empirical studies has not been as strong as might be anticipated (Tsui and Ochoa, 1992; Boulier, 1985).

Although no doubt due in part to measurement difficulties, the failure to observe stronger relationships between measures of physical access and service utilization in prior studies suggests that other factors might also be relevant in defining the concept of access. Indeed, some researchers have argued that access to services is not merely an issue of physical access, but one that involves other dimensions as well (Chayovan et al., 1984; Foreit et al., 1978). Foreit, for example, suggest the following as relevant dimensions or elements of accessibility (the authors used the term "availability" in the original text):

- geographic, or physical, accessibility,
- economic accessibility,
- administrative accessibility, and
- cognitive accessibility.

Working definitions of these elements or dimensions are as follows:

Economic accessibility refers to the extent to which the costs of reaching service delivery or supply points and obtaining contraceptive services/supplies are within the economic means of a large majority of the target population.

Economic barriers affect contraceptive use both by discouraging potential clients from seeking out services and by making contraceptive continuation difficult.

Administrative accessibility refers to the extent to which unnecessary rules and regulations that inhibit contraceptive choice and use have been eliminated; for example, restricted clinic hours for family planning services, restrictions on the distribution of contraceptive during clinic hours for other services (e.g., child immunizations, growth monitoring, etc.) .

Cognitive accessibility concerns the extent to which potential clients are aware of the locations of service/supply points and of the services available at these locations. For example, a client can be unaware of the existence of a SDP, even though it is physically accessible.

To these we add a fifth dimension: psycho–social accessibility, or the extent to which potential clients who desire to space or limit fertility are unconstrained by psychological, attitudinal or social factors in seeking out family planning services. In some settings, for example, potential clients may be fearful of utilizing nearby services because of negative social stigma attached with doing so, may be wary of certain procedures (e.g., pelvic examinations), or may be unable to seek services because women are not permitted to travel alone to obtain contraceptive services or supplies.

Thus, service accessibility is conceptualized as consisting of five elements or dimensions that constitute potential barriers that must be addressed by family planning program management in order to promote wide utilization of available services; that is, to "get clients to the door."

In considering indicators for each of these elements, it should be borne in mind that considerably more work has gone into the measurement of physical access than the other elements of accessibility. As a result, there is a considerably stronger empirical basis for proposing indicators for this element than for the others. Even for this dimension, however, there is no clear consensus as to the most appropriate single indicator (Hermalin and Entwisle, 1988; Tsui and Ochoa, 1992). Accordingly, we have adopted the strategy of proposing a single, "illustrative" indicator for each element or dimension of service accessibility and a list of possible alternatives, with the hope that review and field testing by family planning programs and technical assistance organizations over the next few years will contribute to building consensus around a small set of indicators for each element.

Indicator

NUMBER OF SDPS LOCATED WITHIN A FIXED DISTANCE OR TRAVEL TIME OF A GIVEN LOCATION (I.E., SERVICE DENSITY)

Element

Physical accessibility

Definition

The number of different contraceptive service and method distribution points that are located within a specified distance (e.g., 30 kms) or travel time (e.g., 2 hours) from a given reference location.

Data Requirements

Information on the location of and contraceptive methods available at SDPs within the community.

Data Source(s):

Careful mapping of SDPs (preferred).

Reports by knowledgeable local informants on locations of SDPs (less preferred).

Reports by respondents in surveys on locations of SDPs (least preferred).

Purposes and Issues

This indicator provides a summary measure of the density of family planning services available to clients (and potential clients) within a defined geographic area.

A number of alternative indicators of physical access have been proposed in the research literature, including:

- the proportion of villages or city neighborhoods that have an SDP located in the village or neighborhood;
- the average distance to the nearest SDP, or an index of average distances to a designated set of SDPs that jointly offer the full range of family planning services or methods available in a given country;
- the average travel time to the nearest SDP, or an index of travel times to designated SDPs;
- the average number of hours per month spent obtaining contraceptive services and supplies;
- the number of types of contraceptive methods and services available from any facility/source within the community; and
- the mean number of family planning worker visits received during a specified period (e.g., a quarter) by program clients (for use in household distribution/outreach programs).

While the various indicators tend to be correlated and thus seem to be measuring the intended phenomena to some degree, there are a number of substantively important differences among these indicators. First, it should be noted that some of the indicators pertain to access to family planning services generally, and not to access to particular contraceptive methods or brands. This issue is important, since proximity to SDPs that offer an insufficient range of services and/or methods does not constitute access in the manner intended by the indicator.

Secondly, while some indicators are based upon “objective” measures, others are based upon subjective information provided either by family planning clients (or potential clients) themselves or by knowledgeable local experts. Prior research suggests that information provided by survey respondents may be influenced by their service utilization experience (e.g., users being more knowledgeable than non–users), and thus might provide misleading measures of accessibility (Tsui and Ochoa, 1992).

Indicator

Cost of One Month's Supply of Contraceptives as a Percentage of Monthly Wages

Element

Economic accessibility

Definition

"Costs" for this indicator refer to out–of–pocket expenses for contraceptive supplies and services.

Data Requirements

Information on monthly expenditures on contraceptive supplies and services and estimated monthly income.

Data Source(s)

Population–based surveys' information on service and supply costs–fees may also be available from SDP records.

Purpose and Issues

This indicator provides a measure of the relative economic burden represented by monthly service and supply costs of contraceptive use.

Service and supply costs exceeding one percent of monthly wages for a significant proportion of clients would be considered an indication of the existence of economic barriers to contraceptive use (Ross et al., 1992).

The illustrative indicator for this element was chosen from among several alternatives in large part because the data required for its computation are the most likely to be available in a reasonably large number of developing country settings among the alternatives. However, it should be recognized that the indicator suffers from several important limitations.

One limitation is that it does not take into account other costs of contraceptive use that may be just as or perhaps more important barriers to contraceptive use than direct service or supply costs. For example family planning clients may also incur out–of–pocket expenses for transportation to and from the SDP and (possibly) for child care, as well as "opportunity costs" of time spent travelling to and from the SDP and waiting for service or supplies once reaching the SDP. Thus, a more valid measure of the costs of family planning services would also include these costs in the computation of the indicator.

Another issue concerns the stream of income that should be considered in computing the indicator. Since not all income (gross income) is likely to be available for use in paying for contraceptive services, a more appropriate specification of the indicator would limit the denominator of the measure to monthly *disposable income*. Furthermore, since men and women do not have equal access to household financial resources in many societies, a further refinement might be to limit the denominator of the measure to income or wages *controlled* by the client (especially female clients).

It should be recognized, however, that these refinements add to the data requirements for computing the indicator. In many countries, the required information may be available only from special studies. For most practical purposes, the simpler indicator should suffice to guide program management decisions regarding the affordability of contraceptive services. In programs where cost recovery and sustainability are priority management issues, however, the added costs of gathering data required for the more refined measures may be justified.

Indicator

RESTRICTIVE PROGRAM POLICIES ON CONTRACEPTIVE CHOICE

Element

Administrative accessibility

Definition

Rules and regulations that restrict choice of contraceptive methods for reasons unrelated to medical considerations are mandated by the program; for example, prohibitions against nulliparous women receiving the NORPLANT® implant.

Data Requirements

Eligibility criteria for contraceptive methods offered by the program.

Data Source(s):

Program documents outlining policies and regulations.

Purposes and Issues

This indicator is intended to provide a measure of the existence of barriers to contraceptive choice in the form of unnecessary formal program policies, regulations and procedures; that is, restrictions mandated at the policy/program level beyond those that are justified on medical grounds.

Other examples of the types of barriers that are intended to be measured by the present indicator are:

- requirements for spousal consent for certain contraceptive methods;
- restrictions on certain methods based upon marital status or parity;
- requirements for blood tests or pelvic examinations prior to the distribution of oral contraceptives;
- requirements for multiple visits to receive certain contraceptive methods (IUDs, for example);
- requirements of direct physician involvement in the disbursement of oral contraceptives;
- a required waiting period of several days between counseling for voluntary surgical contraception and the actual procedure; and
- regulatory requirements and procedures that restrict or delay the availability of contraceptive methods that are widely viewed as safe in the scientific community.

The reader is referred to the article by Shelton et al., (1992) for further discussion of these issues.

Indicator

Percentage of the Target Population who Know at Least One Source of Contraceptive Services/Supplies

Element

Cognitive accessibility

Definition

The proportion of the population of reproductive age (total, or by gender) that can name one or more specific location(s)/source(s) where contraceptive methods and/or services may be obtained.

Data Requirements

Responses to survey questions on knowledge of the location of and contraceptive methods supplied by services offered in a given community.

Data Source(s)

Population–based surveys.

Purposes and Issues

This indicator provides a crude measure of the level of knowledge or awareness in the general population as to a source of contraceptive services/supplies. As a measure of cognitive accessibility, it provides an indication of the proportion of the population for which existing services are inaccessible due to lack of knowledge of where they are offered. Although cognitive accessibility is likely to be correlated with physical accessibility, the "cognitive map" of services available may be incomplete or inaccurate, thus acting to limit utilization of services.

In most countries, a sizeable proportion of the population is aware of at least one source of contraceptive services or supplies, but not necessarily of alternative sources of services that may be physically available to them. This lack of awareness may, in some instances, have important implications for contraceptive use. For example, clients or potential clients may be aware only of the nearest facility or supply point offering family planning services, but not of alternative service points to which they have physical access that may offer different methods or perhaps higher quality services. In such cases, dissatisfaction with the range of methods offered or the quality of the services provided at the nearest SDP may result in a decision not to adopt or to discontinue contraception.

Accordingly, a more refined indicator of cognitive accessibility might assess the extent to which the "cognitive map" of services and methods in the population at large is consistent with an "objective map" of service and method availability; that is, the extent of awareness of physically available sources of services. For such purposes, an indicator such as "the proportion of the population who know what methods are available in a defined area and where they can be found" might be considered. This alternative indicator would require conventional survey questions on knowledge/awareness of sources of contraceptive methods and supplies and information from an "objective" mapping of service delivery points within a specified radius.

Indicator

Percentage of Non–use Related to Psycho–social Barriers

Element

Psycho–social accessibility

Definition

The proportion of women desiring to limit or space births who are not using a contraceptive method due to barriers of a psycho–social nature; for example, fears of negative social stigma associated with contraceptive use, fears regarding submitting to specific procedures (e.g., pelvic examinations), fear of side–effects, social restrictions on women traveling alone to seek services, etc.

Data Requirements

Information on reasons for non–use of contraception among women who desire to limit or space births.

Data Source(s):

Population–based surveys or focus group discussions.

Purposes and Issues

This indicator is intended to provide a measure of the extent to which access to otherwise accessible family planning services is limited by barriers of psychological, attitudinal, or social origin. Much has been written on this topic in studies seeking to identify the factors underlying the non–use of contraception among women with an apparent need for family planning, although not necessarily under the heading of psycho–social factors affecting access to services.

In defining the indicator, it is important to distinguish between psycho–social factors affecting demand for family planning and psycho–social barriers to contraceptive use. The former consist of broad social and psychological factors that influence societal family size norms, demand for children (see Chapter V), and the acceptability and social–psychological costs of family planning. The latter, on the other hand, consists of obstacles that must be overcome by individuals who are motivated to obtain family planning services; that is, have a demand or need for family planning (see Chapter VI for further discussion of this concept). For individuals who do not desire to limit or space future births, the present indicator is not relevant since these individuals do not have a need for family planning services. For individuals desiring to control future fertility, however, barriers of psycho–social origin may inhibit their seeking out available services. It is barriers or obstacles of this type that are the focus of the present indicator.

Given the context–specific nature of factors falling under this heading, it is likely that the specific factors playing an important role in the non–use of contraception will vary from setting to setting. Thus, it may be that the exact numerical figure associated with a particular barrier or factor is less important than the rank ordering of problems. In view of this, data from focus groups (that do not provide results in quantitative terms such as percentages or ratios) may be more valuable in identifying barriers of this type than data derived from structured interviews.

QUALITY OF CARE

Illustrative Indicators:

- Number of contraceptive methods available at a specific SDP
- Percentage of counseling sessions with new acceptors in which provider discusses all methods
- Percentage of client visits during which provider demonstrates skill at clinical procedures, including asepsis
- Percentage of clients reporting "sufficient time" with provider
- Percentage of clients informed of timing and sources for resupply/revisit
- Percentage of clients who perceive that hours/days are convenient

Quality of care has become a major issue for practitioners in international family planning programs in recent years. This is not to say that program personnel were not concerned previously with the needs and interests of the client population. However, program managers in many countries around the world are now adopting a more systematic approach to institutionalizing the process of continuous quality improvement in family planning service delivery (CEDPA and FHI, 1992). Concurrently, evaluation specialists have wrestled with the methodological issue of measuring quality.

It is useful to examine the issue of quality in historical perspective. The early family planning initiatives in the 1950s and 1960s were motivated by demographic concerns; the vanguard countries developed family planning programs in an effort to control rapid population growth. As such, the ultimate objective of these programs (and the majority that have followed) was to reduce fertility. This translated to a strong emphasis on the quantitative aspects of service delivery. How many acceptors entered the program each year? What volume of contraception was distributed? What percentage of the population at risk used a contraceptive method? In this sense an interest in quantity rather than quality has predominated in international family planning over the years.

However, there has been a growing recognition of the link between quantity and quality (Jain, 1989; Jain, 1992). Contraceptive prevalence is determined not only by the number of couples who adopt family planning but also by the length of time they continue to use it. Moreover, new acceptors will more readily come forward for a service that they perceive as of good quality than to a substandard one. Thus, even those driven by demographic goals and program targets are coming to recognize the importance of quality in family planning programs. In addition, there are many who support the quality movement on the humanistic premise that all women deserve to receive the best services possible within the limits of local conditions (Jain et al., 1993).

Another possible reason for the emphasis on quantity over quality has been the difficulty of measuring the latter. While it is relatively easy to count the number of contraceptives distributed, measuring quality of care is considerably more complex.

Two events in the late 1980s contributed to advancing the assessment of quality. First, a

conceptual framework for describing quality of care was developed and published by Judith Bruce (1990). This framework includes six elements: choice of methods, information given to clients, technical competence, interpersonal relations, continuity mechanisms, and constellation of services (later modified by some practitioners to "appropriateness and acceptability of service").

Second, a subcommittee on Quality of Care was organized under the USAID Task Force on Improving Family Planning Program Performance Indicators. This group adopted the Bruce framework and provided illustrative indicators for the six elements within the framework. They encouraged colleagues in this field to "let 100 flowers bloom" (i.e., to experiment with different approaches to assessing or measuring quality) (Subcommittee on Quality Indicators in Family Planning Service Delivery, 1990). This effort paved the way for various organizations to develop and test different methods of assessing quality according to their institutional needs.

In early 1992, a Service Delivery Working Group (SDWG) was convened under The EVALUATION Project to look at issues of quantity, quality, and cost within family planning programs. This group identified quality as an area that would most benefit from further refinement of indicators and methods of evaluation. The SDWG as well as a subcommittee of this group worked to define a core list of measurable indicators consistent with the Bruce framework (which has been widely endorsed by the USAID cooperating agencies) that could become a standard across agencies working in this area. This process benefited greatly from the work of organizations actively involved in the measurement of quality, such as The Population Council through its Situation Analysis (Fisher et al., 1992; Mensch et al., 1993).

The core list of indicators developed by the subcommittee in June 1992 is shown in Appendix C (The EVALUATION Project, 1992b). Some have suggested that the current list of 42 indicators should be reduced to a more manageable number, especially in light of the probable intercorrelation among indicators on a given element (i.e., a program that would score well on one measure of method choice would likely score well on a second measure of method choice). However, there has been so little empirical work with these indicators that it is not yet known which are most practical to collect at the field level and which correlate most highly with outcome measures (e.g., continuation, prevalence). Thus, the full list has been maintained during the current testing phase. However, for the purpose of this Handbook, only one indicator per element has been selected and described in detail. As such, the indicators in the text should be viewed as illustrative.

Indicator

Number of Contraceptive Methods Available at a Specific SDP

Element

Choice of methods

Definition

"Number available" refers to those observable at a given SDP, non–expired, for which a trained provider is available to administer (e.g., IUD insertion, tubal ligation).

Data Requirements

A count of different types of methods available at the SDP that are not expired, combined with verification that appropriate service providers are available to deliver the methods to clients.

Data Source(s)

Observation at the SDP (e.g., as is done in situation analysis).

Purpose and Issues

This indicator measures the "necessary but not sufficient" condition for ensuring a full range of contraceptive methods for clients, one of the six elements in the Bruce framework.

For a method to be available, action is required at two levels: (1) at the managerial level, to ensure that non–expired contraceptive supplies and trained providers are available at a given point; and (2) at the provider level, to effectively offer all appropriate methods to a given client. (In this sense, "appropriate" is used in reference to the client's reproductive intentions and possible medical contraindications.) The observation of existing supplies measures the manager's ability to stock all nationally–approved methods for the type of facility in question. For a measure of the provider's willingness to offer all appropriate methods, see Indicator #6 in Appendix C.

Indicator

Percentage of Counseling Sessions with New Acceptors in Which Provider Discusses All Methods

Element

Information given to clients

Definition

"All methods" refers to those that are available at the SDP and appropriate to the specific client (in terms of reproductive intentions and possible contraindications).

Data Requirements

Checklist of the methods described by the provider.

Data Source(s)

Observation of provider–client transaction (by an observer, or if possible, by a mystery client); exit interview (if observation not possible).

Purpose and Issues

This element refers to the question: is the information provided to clients complete, accurate, and understandable? It refers to the content of what is said (in contrast to Element #4 that reflects the affective aspect of how it is said).

If the provider is aware of being observed, he/she may give a more thorough presentation of the methods than would normally be the case. Thus, the use of a mystery client (member of research team who presents herself as a client for the purposes of observing service delivery "under normal conditions") is particularly appropriate for obtaining this information. However, this technique will not work well in rural settings or SDPs with a known clientele (e.g., CBD program) where any "new face" might be viewed with suspicion.

One could also question clients as they were leaving the facility as to the methods that were discussed in the session. This approach, however, is subject to recall bias on the one hand and to the mention of methods known beforehand (that were not in fact presented in the session) on the other.

Indicator

Percentage of Client Visits at Which Provider Shows Skill at Clinical Procedures, Including Asepsis

Element

Technical competence

Definition

The assessment of "skill" should be based on adherence to guidelines for service delivery established at the national and/or institutional level. The specific "clinical procedures" to be included must be identified by those doing the assessment.

Data Requirements

- List of procedures to be assessed
- List of criteria for assessing competence
- Score on each procedure

Data Source(s)

Observation by team of clinicians judged to be expert in this field.

Purpose and Issues

Technical competence is critical to the quality of care issue. Although the client may not be able to judge whether the correct clinical procedures are used during a given clinic visit, it is incumbent on the family planning program to establish standards for the delivery of family planning services and to monitor compliance with those standards. No matter how modest the physical facilities, a minimum level of technical competence by service providers is imperative.

This vigilance can be carried to the extreme that it becomes a medical barrier if unnecessary tests and procedures are required in the name of safeguarding the health of the patient. However, this fact does not diminish the importance of requiring service providers to be technically competent.

This "indicator" requires not one but many points of measurement, corresponding to the different procedures required to provide clients with the full range of contraceptive methods. An assessment of this type could be extensive, covering many diverse aspects of service delivery, or it could be limited to the 2–3 aspects judged by clinical experts to be most essential in safeguarding clients' health.

Whereas this type of assessment is commonplace in medical facilities, it can be adapted to fit CBD and contraceptive social marketing programs as well. Specifically, one would identify what a CBD worker or pharmacist/store clerk should do with respect to each method offered to clients, and one would proceed to design an instrument that would allow for a scoring of these workers in a field situation. However, since these types of workers do not perform clinical procedures, this assessment will focus mainly on the counseling aspects of service delivery.

Indicator

Percentage of Clients Reporting "Sufficient Time" With Provider

Element

Interpersonal relations

Definition

"Sufficient" is defined by the client according to his/her own expectations and perceptions; based on number of clients questioned during a given period (e.g., day or week).

Data Requirements

Response from client to question on questionnaire.

Data Source(s)

Exit interview, follow–up survey in client's home.

Purpose and Issues

This indicator forms part of a set of indicators that measure the clients' attitudes toward their interactions with service providers. Such attitudes are hypothesized to contribute to the client's overall satisfaction with the services. (Other items in this block of indicators concern total time of the visit, ease/difficulty in asking questions, courtesy and respectful treatment by staff.)

This set of items is highly susceptible to a courtesy bias, especially if the client believes the interviewer works for the family planning program. Moreover, in societies where low income populations feel they must accept whatever quality of services are available, respondents may not fully comprehend the nature of the question being asked; they may not realize that better services could be offered; or they may feel it is inappropriate to make negative comments about the service providers or the services offered (Subcommittee on Quality Indicators in Family Planning Service Delivery, 1990).

Given the above limitations, one might question the value of obtaining respondents' attitudes on these issues. However, those working in quality of care would argue that for decades, donors and program managers have decided how services should work, failing to take into consideration the client perspective. One could obtain a much more reliable measure of "time spent with service provider" by using an outside observer to clock elapsed time; yet one would not know how the client reacted to this wait. A thirty minute delay might seem long in one society, short in another. Thus, what is important is not the actual elapsed time but rather the client's attitude toward it, which is expected to affect his/her probability of adopting and sustaining contraceptive use.

Indicator

Percentage of Clients Informed of Timing and Sources For Resupply/Revisit

Element

Mechanisms to ensure continuity

Definition

Clients need to know:

- when they need to return;
- where to return (if other than same SDP); and
- where to obtain resupply (if other than same SDP, especially if alternative sites are more convenient).

Data Requirements

Evidence that provider informs client of timing and sources for resupply/ revisit.

Data Source(s)

Observation (preferably with a mystery client); exit interviews with clients.

Purpose and Issues

The "need to return" will depend upon the personal characteristics of the client, as well as the type of method prescribed. Thus, this part of the evaluation should measure the existence of clearly stated criteria with regard to the timing of revisits and provider compliance with these criteria.

Indicator

Percentage of Clients That Perceive That Hours/Days Are Convenient

Element

Appropriateness and acceptability of services

Definition

(self–explanatory)

Data Requirements

Responses from clients in structured interviews or focus groups.

Data Source(s)

Exit interview; follow–up of client; focus group of clients.

Purpose and Issues

The indicators under this element cover a number of different aspects of service provision that affect client satisfaction. Additional indicators measuring appropriateness and acceptability of services include the client's perceptions with respect to privacy for counseling, privacy for the physical exam, waiting time, amount of time with the provider, the socio–demographic characteristics of the staff (gender, ethnic group, age), and adequacy of the facility. Thus, while somewhat of a catch–all category, these issues strongly influence a potential client's decision to seek services, adopt a contraceptive method, and sustain use.

In societies where low income populations have low expectations of the health system, clients may feel they should take what they can get; similarly, the hours/days may be so fixed in a given setting that none would bother to question them. And as mentioned in connection with the earlier indicator "client reports sufficient time with provider," there may be a strong courtesy bias, especially if the interviewer is perceived to work for the family planning program.

PROGRAM IMAGE

- Number and type of activities to improve the public image of family planning during a reference period (e.g., one year)
- Percentage of target population favorable to the (national) family planning program

Of the three components of service outputs—accessibility, quality, and program image—far more work has been done on the first two than on the third. Yet based on the experience of marketing specialists in the private sector, the image of the program can have a marked effect on public reaction to the philosophy of family planning and individual attitudes toward contraceptive use.

There are numerous examples of efforts to shape/improve the public image of a program, often linked to structural changes within the organization. In Indonesia, the Blue Circle Campaign was launched to create a strong image of quality in the government's family planning service provision. In Guatemala, APROFAM (the private family planning association) came under unrelenting attack in the media from conservative factions in the mid-1980s. It responded by broadening its services to include a wider range of reproductive health interventions and positioning itself as a major contributor to improved health among women throughout the country.

How successful are such efforts? Given the cost involved in tracking public attitudes toward family planning, efforts to improve program image have often been launched without rigorous monitoring. What limited research funds exist are often directed to audience research on the most appropriate audiences and message, or on pretesting of materials in pre-production form.

In the past, one common source of data on public opinion toward family planning was the traditional Knowledge-Attitude-Practice (KAP) survey. However, as this type of survey evolved into a Contraceptive Prevalence Survey (CPS) and subsequently into the DHS, the "soft" question on general attitudes toward the acceptability of family planning was dropped in moderate to high prevalence countries on the grounds that responses to such questions were of little value where a sizeable proportion of respondents were current contraceptive users. At present, the use of this type of question is limited largely to surveys intended to monitor a specific communication effort and in low prevalence DHS countries.

In this section we have equated attitudes toward family planning in general with attitudes toward the (national) family planning program. Whereas there may be cases where the public clearly distinguishes between the two, in general the two tend to blur together. If the distinction were important, one could modify the proposed indicators accordingly.

At this time, only two indicators are proposed for program image.

Indicator

Number and Type of Activities to Improve the Public Image of Family Planning During a Reference Period

Definition

Quantity of communications (either via mass media or interpersonal channels) or other activities that explicitly try to cast the program's goals and objectives in a favorable light for the purpose of enhancing institutional image and/or improving public attitudes toward family planning in general.

Data Requirements

Inventory of communications disseminated during a given period (e.g., one year).

Data Source(s)

I–E–C department, advertising agency contracted for this purpose.

Purpose and Issues

While this indicator can theoretically provide this type of information, most evaluators would prefer to know the effectiveness of such communications programs in changing attitudes (assuming this as one of the objectives). Thus, the following indicator on public attitude toward family planning is by far the more practical and relevant.

Indicator

Percentage of Target Population Favorable to the (National) Family Planning Program

Definition

Percentage of population who give a positive response regarding attitude toward family planning in general or toward the national program in particular.

Data Requirements

Percentage giving positive response to attitudinal question on survey.

Data Source(s)

Population-based survey of target population.

Purpose and Issues

This indicator can be used with respect to family planning in general (encompassing different providers of family planning services in that country) or to a specific program or institution (e.g., the social marketing program). It serves to identify public sentiment that may influence individual behavior, and it allows for monitoring the legitimization of family planning over time.

This indicator may be of particular use among hard-to-reach populations (often on a project level). For example, qualitative research among Mayan Indians in one *departamento* of Guatemala indicated that many people equated family planning with killing the children yet to be born to a particular woman (Ward et al., 1992). In this situation, it would be useful to monitor the effect of program interventions on this attitude over time.

Chapter V

Fertility Demand

- Mean desired family size
- Desire for additional children
- Wanted status of previous births
- Wanted total fertility rate (WTFR)

Chapter V

INDICATORS TO MEASURE DEMAND FOR CHILDREN (FERTILITY DEMAND)

The conceptualization and measurement of demand for children has been the subject of considerable research. Based on micro–economic and consumer–demand theories of fertility decision–making, demand for children is usually taken to refer to the number of children parents would choose to have if there were no subjective or economic problems involved in regulating fertility (Bulatao and Lee, 1983; Easterlin, 1978; McClelland, 1983; Espenshade, 1977). Demand is seen as a matter of relative preference for children versus other consumption activities, and it includes preferences for the timing, spacing, and gender of children in addition to quantity (Bulatao, 1981). Some researchers would add the condition that the concept of demand must also have a time reference (e.g., at the time of a survey), as it is assumed that parents gain additional and more accurate information on the costs and benefits of children as they pass through the reproductive years, and it is thus possible that demand may change over time (Lee, 1980; Pullum, 1980).

As noted in the introduction to this Handbook, demand for children is influenced by a large number of societal– and individual–level factors. An important feature of The EVALUATION Project conceptual framework is that fertility demand is viewed as also influenced by characteristics of family planning programs. Specifically, features of the family planning supply environment (indicators of which are presented in Chapters II and III of this Handbook) are considered relevant in influencing the demand for children and the demand for family planning services (indicators of which are covered in Chapter VI), should women or couples decide to take action to implement their fertility preferences.

For program management purposes, indicators of fertility demand are useful in several respects. First, they provide information on prevailing societal norms and preferences regarding family size as well as changes in these over time. Given strong empirical evidence linking fertility preference measures to current and future contraceptive use and fertility levels (Westoff, 1990; Bongaarts, 1991a), indicators of fertility demand may provide valuable information on the level of future demand for family planning program services.

Secondly, when combined with information on current fecundity and contraceptive use, indicators of fertility demand provide a basis for deriving two useful program outcome indicators: (1) the level of "unmet need' for family planning services, and (2) the proportion of total demand for family planning services being satisfied by current contraceptive use (see Chapter VI for discussions of these indicators).

There has been a lively debate over the years among population researchers as to whether the concept of demand can be validly measured through conventional survey questions on fertility preferences and intentions (Demeny, 1988; Hauser, 1967; McClelland, 1983). Methodological research suggests that "standard" survey questions on fertility preferences and intentions do appear to effectively "tap" the concept of demand in (roughly) the economic sense of the concept, although some lines of questioning appear more valid than others (Bulatao and Lee, 1983; McClelland, 1983). Furthermore, given a rather large body of empirical evidence showing a high degree of consistency between responses to standard survey questions and actual contraceptive and fertility behavior in a large number of developing country settings, apparently meaningful measures of fertility aspirations can indeed be derived from the types of questions typically included in demographic surveys (Westoff, 1990; Bongaarts, 1991a).

Four indicators, with strengths and limitations noted in the discussion of each, are proposed in this chapter as relevant measures of fertility demand.

Indicator

MEAN DESIRED FAMILY SIZE

Definition

The average number of children that women (or couples) of reproductive age would choose to have if they could have exactly the number desired.

Data Requirements

Responses to questions on desired number of children (see below for details on the wording of questions).

Data Source(s)

Population–based surveys or facility–based data.

Purposes and Issues

Desired, or ideal, family size is the best known and most widely available indicator of fertility preferences or demand. Most if not all fertility/family planning surveys conducted in recent years include the required question(s) for the indicator. As is evident from the definition above, the indicator provides a measure of the level of completed fertility desired by women or couples under the idealized circumstances that they are able to perfectly control their fertility and have exactly the number of children desired.

The most common (and preferred) source of data for the indicator is population–based surveys. The DHS asks separate questions of women with and without living children. Women with no surviving children are asked, "If you could choose exactly the number of children to have in your whole life, how many would that be?" Women with surviving children are asked, "If you could go back to the time when you did not have any children and could choose exactly the number of children to have in your whole life, how many would that be?" Based on this line of questioning, the terms "desired" and "ideal" family size may be, and are in practice, used interchangeably. Roughly comparable questions were included in the earlier World Fertility Survey (WFS) and Contraceptive Prevalence Survey (CPS) programs.

Information on reproductive preferences is also available from client record systems in clinic–based family planning programs in many settings. The information is usually obtained, along with a client's reproductive history, during an initial interview for new acceptors. Two important differences between these data and survey–based data should be noted. First, clinic–based data normally pertain to desired family size given client's current number of children, while survey data represent a more idealized measure; that is, the number of children women would have if they could "start over." Secondly, it should be recognized that data on desired family size obtained from facility–based sources reflect the demand for children among women who seek clinic–based family planning services and are not necessarily indicative of the level of demand for children in the larger population. For this reason, data from population–based surveys are normally preferred as the source of information on demand for children at the population level.

While the indicator is widely used, methodological research suggests that it tends not to be an unbiased indicator of demand, and it is thought to suffer from the following biases (Westoff, 1991; Bongaarts, 1990):

- the tendency for respondents to give "normative" responses;
- the tendency of high–parity women to rationalize unwanted pregnancies by reporting desired family sizes that are equal to or exceed their current parity (that is, ex–post facto rationalization); and
- the inability or unwillingness on the part of respondents to quantify their fertility desires (for example, "as many as possible" or "up to God").

Research evidence to date suggests that other indicators considered in this chapter may provide more valid measures of the level of demand for children.

Indicator

DESIRE FOR ADDITIONAL CHILDREN

Definition

The number or proportion of women (or couples) of reproductive age who want to have a (another) child or, conversely, desire not to have additional children.

Data Requirements

Numbers or proportions of respondents reporting that additional children are/are not desired.

Data Source(s)

Population–based surveys or facility–based data.

Purposes and Issues

This indicator is widely used in surveys to identify women (or couples) with a demand for additional children on the one hand and those who do not desire additional children and thus have an apparent need/demand for fertility limitation on the other. In the DHS, non–pregnant women married or in union are asked, "Would you like to have a (another) child or would you prefer to not to have any (more) children?" Women who are pregnant (or uncertain of their status) at the time of the survey are asked, "After the child you are expecting, would you like to have another child or would you prefer not to have any more children?"

On the basis of responses to these questions, respondents may be divided into two categories: those with demand for additional children and those desiring to terminate childbearing, with women in the latter category considered as having a "demand for family planning" (see Chapter VI). A recent publication illustrates the use of DHS–like questions in examining cross–national differentials and trends in intentions to terminate fertility (Westoff, 1991).

Responses to this type of question may also be used in conjunction with information on current fecundity and contraceptive use in assessing the level of unmet need for family planning (see Chapter VI).

Comparable information may sometimes be available from service statistics of clinic–based family planning programs. Questions similar to those included in the DHS are often asked of (at minimum) new clients in order to determine the appropriateness of different contraceptive methods in relation to reproductive intentions: that is, methods appropriate for limiting versus spacing. The caveats regarding the use of facility–based data noted on the previous indicator also apply to facility–based data on this indicator.

Despite earlier concerns as to validity of survey questions of this type in predicting actual fertility behavior, recent studies have provided rather convincing evidence of strong aggregate–level associations between expressed desires for additional children on the one hand and patterns of current contraceptive use and current and future fertility on the other (Bongaarts, 1990; Westoff, 1991). The indicator is currently viewed as being relatively unbiased, as there are no obvious reasons for respondents to over– or under–report preferences to continue childbearing.

Indicator

WANTED STATUS OF PREVIOUS BIRTHS

Definition

The number or proportion of births occurring during a specified prior period of time that were "wanted," or conversely, were "unwanted."

Births are classified as wanted when respondents report having desired additional children at the time of becoming pregnant with the reference birth.

Unwanted births are those for which respondents report having not desired additional children at the time of becoming pregnant.

Data Requirements

Responses to retrospective questions on whether or not respondents had desired additional children at the time of becoming pregnant during a specified interval of time (for example, for births occurring in the 2–5 years prior to a survey).

Data Source(s)

Population–based surveys.

Purposes and Issues

This indicator has been proposed as another alternative to the desired family size measure of fertility demand. In principle, the indicator provides a simple and relatively direct measure of "wanted fertility," a measure of demand for recent fertility. Unlike the two previous indicators of demand, this indicator attempts to measure demand at a specified point in the past (at the time of the last pregnancy) instead of on the basis of current reproductive intentions.

Wanted births consist of births to women who desired a (another) child at the time of becoming pregnant with the reference child, plus births resulting from pregnancies that were desired, but not at the time that they occurred (that is, timing failures). Unwanted births are those occurring to women not desiring additional children at the time of becoming pregnant.

The indicator may be derived from a survey question on whether or not respondents desired to have additional children at the time of becoming pregnant for births occurring in the period just prior to the survey. In the DHS, for example, the following question is asked regarding all births in the five years prior to the survey: "Just before you became pregnant with (name of child) did you want to have more children then, did you want to wait longer, or did you want no more children?" Desired births consist of those in the first two categories.

The proportion of previous births that are reported as not desired provides a conceptually simple indicator of the extent of fertility control failures. Due to potential methodological problems, however (see details below), this indicator has not been widely used. Researchers have long felt that the indicator may be seriously biased towards overstating the actual level of "desired" fertility due to reluctance on the part of survey respondents to admit to unwanted pregnancies in survey interview situations. A recent study of seven DHS countries provides empirical support for this view (Bongaarts, 1990). In the seven countries, the total fertility rate based upon reports of the wanted status of births in the five years preceding the survey (referred to as "the reported wanted fertility rate") was on average 12 percent higher than the desired fertility rate, which itself is thought to overstate the "true" level of wanted fertility.

Despite these reservations, a recent study using DHS data from Indonesia concluded that the level of unwanted fertility implied by the indicator, estimated at 28 percent of births in the five years prior to the survey, was plausible given its consistency with a number of behavioral variables, suggesting that the indicator might be sufficiently unbiased to be useful in some settings (perhaps in settings where fertility control is widely practiced and respondents are more forthcoming about fertility control failures in survey interview situations) (Weller et al., 1981).

Indicator

WANTED TOTAL FERTILITY RATE (WTFR)

Definition

The number of children that would be born per woman (or per 1,000 women) if she/they were to pass through the reproductive years bearing children according to a current schedule of age–specific fertility rates if only "desired" or "wanted" births occurred.

For this indicator, "wanted" births are defined taking into account both desired family size (as defined earlier in this chapter) and the number of surviving children. All births during a specified reference period (usually the 2–5 years prior to a survey) that do not cause a respondent's number of surviving children to exceed the stated desired family size are classified as wanted. Births that cause the number of surviving children to exceed the desired family size are considered unwanted.

The indicator is calculated as follows:

$$WTFR = 5 \sum_a (WB_a / E_a)$$

Where:

WB_a = the number of births to women in age group *a* in a given year or reference period that are "wanted,"

E_a = the number of person–years of exposure in age group *a* during the reference period.

Illustrative Computation

Estimate of the WTFR for women aged 15–44 years in the 36 months prior to the survey, Northeast Brazil.

Age Group	Total Births B_a	Wanted Births WB_a	Person–Years of Exposure E_a	Rate/ Woman
15–19	289	260	3820	.068
20–24	619	466	3200	.146
25–29	459	281	2733	.103
30–34	349	151	2324	.065
35–39	218	78	2267	.034
40–44	71	20	1872	.011

(Source of data: Northeast Brazil Demographic and Health Survey, 1991.)

WTFR = 5 (.068 + .146 + .103 + .065 + .034 + .011) = 2.13

Data Requirements

Responses to survey questions on:

- numbers and dates of births during a recent period (typically the 2–5 years prior to a survey);
- desired number of children or family size; and
- number of children ever born and number surviving.

Data Source(s)

Population–based surveys.

Purposes and Issues

The WTFR is a measure of "wanted" fertility, a hypothetical measure of what the total fertility rate (TFR) would be given age–specific fertility rates for a recent past period under the condition that all women's fertility preferences were perfectly realized; that is, if only "wanted" births occurred. The measure represents an attempt to avoid the suspected bias in the wanted status of recent births indicator by defining wanted or desired status on the basis of the consistency (or lack thereof) between the reported desired family size and the number of surviving children, as opposed to on the basis of retrospective reports of fertility intentions at the time of becoming pregnant.

The indicator is calculated as the sum of age–specific fertility rates, or the total fertility rate, after the deletion of births occurring during a specified reference period that cause the number of surviving children of sample respondents to exceed their stated desired number of children.

In the DHS, numbers of births during the specified reference period are derived from the birth history portion of the survey interview, numbers of surviving children from questions on lifetime fertility and survival status, and information on desired family size from survey questions described earlier in this chapter.

The above definition of the WTFR is based upon the work of Lightbourne (1985, 1987) and Westoff (1991) (who labels the measure the "desired total fertility rate" or DTFR). In a recent publication,

Bongaarts (1990) proposes a modified definition of the WTFR in which wanted births are defined on the basis of whether survey respondents desired additional births at the time of a survey instead of on the basis of the comparison of the desired number of children and the number of surviving children. Under this definition, births within a specified reference period are classified as wanted if the respondent reported wanting additional children at the time of a survey.

The argument for the alternative definition is that it is based upon responses to questions on preferences for additional children, an indicator of demand that is thought to be less affected by reporting biases than the desired family size indicator (Bongaarts, 1990). The reader is referred to the earlier discussion of these indicators. Comparison of estimates of the two versions of the WTFR for 48 DHS countries indicates that the two measures are reasonably close for most countries, with an average difference between the measures of about 9 percent – 4.09 versus 3.76 (Bongaarts, 1990). On the basis of available evidence, either version of the WTFR appears to be preferable to using the wanted status of previous births in defining wanted fertility.

The comparison of the WTFR with the TFR provides a relatively straightforward indication of the extent to which observed fertility exceeds desired or wanted fertility. This type of comparison provides program managers and policy–makers with some insight into the potential short– to medium–term demand for family planning services and the potential for fertility decline in the future (Westoff, 1991). In the case of Northeast Brazil, for example, the comparison of the TFR (3.66) with the WTFR (2.13) suggests that a considerable share of current fertility is unwanted and that sufficient latent demand exists in this population; thus increases in contraceptive prevalence and a decline in fertility might be reasonably expected, given the availability and adequate quality of family planning services.

Chapter VI

Family Planning Demand

- Demand for limiting
- Demand for spacing
- Total demand (for family planning)
- Unmet need for family planning
- Satisfaction of demand for family planning

Chapter VI

INDICATORS TO MEASURE DEMAND FOR FAMILY PLANNING

Demand for family planning refers to the desire or motivation of women or couples to control future fertility. As indicated in the conceptual framework sketched in Chapter I, demand for family planning is directly influenced by the (potential) supply of children; that is, the number of surviving children a woman or couple would have in the absence of parity-specific fertility regulating behavior (Bongaarts and Menken, 1983; Knodel, 1983), and a woman's or couple's demand for children, definitions and indicators of which were presented in Chapter V. Demand for family planning is said to exist when the supply of children exceeds the desired or preferred number. Women or couples are assumed to continue (or at least intend to continue) to bear children until such time as the desired number of children is reached, at which time it is assumed that they will be motivated to control future pregnancies in some fashion (assuming, of course, that the desired number of children remains fixed) (Hermalin, 1983).

Given a desire to limit or space future childbearing, whether or not contraception is actually adopted is assumed to be determined by the outcome of an assessment of the economic, social, and psychic costs of fertility regulation and the strength of motivation to control additional pregnancies. Where the demand to regulate fertility exceeds the actual or perceived costs of regulation, it is anticipated that women or couples will choose to adopt contraception.

Indicators of demand for family planning play an important role in both the management and evaluation of family planning programs. Clearly a central function of program management is to provide sufficient resources to satisfy existing and anticipated future demand for family planning services. Information on current and projected levels and structure of demand (that is, the distribution of women or couples with a demand for services by such factors as geography, socioeconomic status, and demographic characteristics) are thus crucial to effective program planning.

Demand for family planning may be divided into two components: demand for limiting and demand for spacing. Women or couples who desire to terminate childbearing are said to have a demand for limiting, while those who wish to postpone future births (but not to terminate childbearing) are said to have a demand for spacing. Since the services and contraceptive methods required by clients who desire to limit as opposed to space further births are different, measures of demand for limiting and spacing provide valuable information on current and projected future program needs in terms of the mix of services and contraceptive methods.

In addition to satisfying existing demand for family planning services, program administrators must also be concerned with two additional program aspects related to the demand for services. A particularly important function of management is the conversion of latent demand into contraceptive use. In many countries of the developing world, there are growing numbers of women and couples whose number of surviving children exceed their expressed number of desired children (Bongaarts, 1991b; Westoff and Ochoa, 1991). In a number of these countries, however, this apparent or latent demand for fertility regulation has yet to result in women or couples seeking family planning services in significant numbers. In some cases, the supply of services exceeds the actual demand for services by a considerable margin. The motivation of women or couples with an apparent demand for family planning to seek program services through I–E–C efforts (see Chapter III) and/or through other improvements in the family planning supply environment (see Chapter IV) are among the central tasks of program management.

Program management must also be concerned with demand generation; that is, with changing perceptions in the target population for family planning services such that the social, economic, and health-related benefits of fertility regulation, as well as the costs and potential health risks of contraceptive use, become more widely understood and accepted.

For both of the above purposes, indicators of the level, demographic, socioeconomic and geographic distribution of demand provide valuable information for use in designing and implementing efforts to influence the level and strength of demand for program services.

From the standpoint of program evaluation, indicators of the level of demand for family planning provide a basis for the assessment of program performance in a number of areas, including: (a) the extent to which existing demand is being satisfied (see the "unmet need for family planning" and "proportion of demand satisfied by contraceptive use" indicators presented in this chapter); (b) the extent to which programs are strategically oriented given the current and anticipated future level and structure of demand; and (c) as a measure of program success in generating demand.

Finally, it should be noted that the calculation of measures of demand for family planning from survey data is rather complex. This is due to the fact that a number of factors have to be taken into account simultaneously in classifying respondents with respect to level of demand. In view of this, the reader will find in Appendix E illustrative tabulations showing the distribution of women from DHS I data from selected countries by various categories important to the calculation of indicators of demand for family planning, as well as a diagram describing how women in different categories are combined in deriving the measures reviewed in this chapter. These illustrative tabulations are intended to assist readers unfamiliar with this material in understanding the logic of the indicators, as well as in performing the calculations on their own.

Indicator

DEMAND FOR LIMITING

Definition

The number or proportion of women currently married or in union who are fecund and who desire not to have additional children.

The indicator is calculated as follows:

$$D_L = C_L + U_L + F_L$$

Where:

D_L = the number or proportion of women currently married or in union with a demand for limiting,

C_L = the number of women currently married or in union desiring no additional children who are currently using a contraceptive method (i.e., met demand),

U_L = the number of fecund women currently married or in union who desire no additional children but are not currently using a contraceptive method, plus the number of currently pregnant or amenorrheic women currently married or in union whose current/last pregnancy was unwanted and occurred while not using a contraceptive method (i.e., unmet demand), and

F_L = the number of currently pregnant or amenorrheic women married or in union whose current/last pregnancy resulted from contraceptive failure (a special category explanation).

Illustrative Computation

Estimated total demand for limiting in Egypt, 1988 (expressed as a percentage of currently married women).

$$D_L = C_L + U_L + F_L$$
$$= 31.9 + 15.0 + 1.4$$
$$= 48.3$$

Source of data: Egypt Demographic and Health Survey, 1988; calculations from Westoff and Ochoa (1991).

Data Requirements

Responses to survey questions on:

- desire for additional children;
- current contraceptive use status;
- current fecundity, pregnancy, and amenorrhea status for women not currently using a contraceptive method;
- the wanted status (with respect to number) of the current/last pregnancy for women currently pregnant or amenorrheic; and
- whether a contraceptive method was being used at the time of the current/last pregnancy among currently pregnant or amenorrheic women (i.e., whether the last/current pregnancy resulted from contraceptive failure).

Data Source(s)

Population-based surveys.

Purposes and Issues

This indicator provides an estimate of the total number of clients who would have to be served by family planning programs if all women who desire to terminate childbearing were to seek family planning services. The estimate of total demand consists of two principal components: (a) the share of total demand for limiting that is being satisfied through current contraceptive use; and (b) unmet need for limiting (see the discussion of the unmet need for family planning later in this chapter for further details on the definition and computation of this indicator).

Following the procedure proposed by Westoff and Ochoa (1991), women who are pregnant or amenorrheic at the time of measurement of the indicator (e.g., at the time of a survey) whose last or current pregnancy resulted from contraceptive failure are also considered to have a demand for family planning, even though they are not currently at risk of pregnancy. The rationale for this is that since the current/last pregnancy for these women was undesired and they had attempted to avoid pregnancy through contraceptive use, they are highly likely to resume contraceptive use at some point in the near future (although perhaps

with a different method) and should thus be included in the estimate of total demand (for limiting).

In order to account for these women in the computation of demand, however, it is necessary to add a term to the equation (F_L) as shown above, since women in this group are neither currently using a contraceptive method nor fecund, which would thus preclude them from being considered in the computation.

For the purposes of this indicator, sub-fecund women are defined as nonpregnant women in union for at least five years who have not used contraception and who have not been fertile and/or have not menstruated in the twelve weeks prior to the survey (Westoff and Ochoa, 1991).

An alternative method of calculating total demand for limiting has been proposed by Bongaarts (1991b), based upon the alternate method of calculating unmet need for family planning described later in this chapter. This approach differs from that described above in the method used in calculating unmet need. Since, as is indicated later in this chapter, the two methods of calculating unmet need produce similar results, estimates of the level of demand for limiting from the two methods will tend to produce roughly similar results.

Although women who are not currently married or in union constitute a sizeable fraction of the women at risk of pregnancy in some societies, measures of demand for family planning are normally restricted to women currently married or in union. The principal rationale for this is to attempt to limit the measure to women at risk of pregnancy. There is also a question about the validity of responses to questions about childbearing intentions among single women (Westoff and Ochoa, 1991). This restriction should, however, be borne in mind when interpreting estimates of demand in societies in which a significant share of childbearing occurs outside of formal marriages and/or unions.

In populations where this restricted definition of pregnancy risk is clearly inappropriate, an alternative approach might be to include all women of reproductive age under a certain age (e.g., age 25) who had ever had sex based upon sexual activity questions such as are found in the DHS in the computation of demand.

Finally, it should be noted that estimates of proportions of women or couples of reproductive age or currently married or in union with demand for limiting are directly obtainable from survey data such as that provided by the DHS. If absolute numbers of women or couples with demand for limiting are desired (for use in projecting future staff and contraceptive supply needs, for example), recent census figures or population projections will be needed in order to convert the survey estimates of proportions into absolute numbers or population totals. This caveat applies to all indicators discussed in this chapter.

Indicator

DEMAND FOR SPACING

Definition

The number or proportion of women currently married or in union who are fecund and who desire to delay the birth of their next child for a specified length of time (for example, for two years from the date of a survey).

The indicator is calculated as follows:

$$D_s = C_s + U_s + F_s$$

Where:

D_s = the number or proportion of women currently married or in union with a demand for spacing,

C_s = the number of women currently married or in union desiring to delay their next pregnancy for a specified length of time who are currently using a contraceptive method (i.e., met need or demand),

U_s = the number of fecund women currently married or in union who desire to delay their next pregnancy for a specified length of time but are not currently using a contraceptive method, plus the number of currently pregnant or amenorrheic women married or in union whose current/last pregnancy occurred earlier than desired and who were not using a contraceptive method at the time of pregnancy (unmet need or demand), and

F_s = the number or proportion of currently pregnant or amenorrheic women married or in union whose current/last pregnancy occurred earlier than desired as a result of contraceptive failure.

Illustrative Computation

Estimated total demand for spacing in Egypt, 1988 (expressed as a percentage of currently married women).

$$D_S = C_S + U_S + F_S$$
$$= 5.9 + 10.1 + 0.5$$
$$= 16.5$$

Source: Egypt Demographic and Health Survey, 1988; calculations from Westoff and Ochoa (1991).

Data Requirements

Responses to survey questions on:

- desire for additional children and, among women desiring additional children, the preferred length of birth interval;
- current contraceptive use status;
- current fecundity, pregnancy, and amenorrhea status for women not currently using a contraceptive method;
- the wanted status (with respect to timing) of the current/last pregnancy for women currently pregnant or amenorrheic; and
- whether a contraceptive method was being used at the time of the current/last pregnancy among currently pregnant or amenorrheic women.

Data Source(s)

Population-based surveys.

Purposes and Issues

This indicator provides an estimate of the total number of clients that would have to be served if all women who desire to space subsequent births were to seek family planning services. As with the demand for limiting indicator, the estimate of total demand for spacing consists of two components: (a) the share of total demand for spacing that is being met through current contraceptive use, and (b) the unmet need for spacing.

Again following the procedure proposed by Westoff and Ochoa (1991), currently pregnant or amenorrheic women whose last/current pregnancy was mistimed due to contraceptive failure are considered to have a demand for spacing even though these women are not at risk of pregnancy. The rationale for including these women in the computation of demand for spacing is the same as described in connection with the demand for limiting indicator. Since contraception was used to attempt to space the current/last pregnancy and women in this category intended to space their next birth, it is highly likely that they will resume contraceptive use at some point in the near future (although perhaps with a different method) and should thus be included in the estimate of

demand. As with the demand for limiting indicator, it is necessary to add a term to the equation as shown previously in order to account for these women in the computation of demand for spacing, since they would not otherwise enter into the computation.

The alternative procedure for computing unmet need for spacing proposed by Bongaarts (1991b) may also be used in calculating the present indicator.

The reader is referred to the observations made on the measurement of demand in populations where childbearing is not confined to marriage in connection with the demand for limiting indicator, as these apply to the present indicator as well.

Indicator

TOTAL DEMAND (FOR FAMILY PLANNING)

Definition

The number or proportion of women currently married or in union who are fecund and who desire to either terminate childbearing or to postpone their next birth for a specified length of time.

The total number of women with a demand for family planning, or total demand, is equal to the sum of the number of women with a demand for limiting plus the number of women with demand for spacing, as defined in the previous indicators.

Total demand is calculated as:

$$D = D_L + D_S$$

Where:

D = the number or proportion of women currently married or in union with a demand for family planning,

D_L = the number of women currently married or in union with a demand for limiting, and

D_S = the number of women currently married or in union with a demand for spacing.

Illustrative Computation

Estimated total demand for spacing in Egypt, 1988 (expressed as a percentage of currently married women).

$$D = D_L + D_S$$
$$= 48.3 + 16.5$$
$$= 64.8$$

Source of data: Egypt Demographic and Health Survey, 1988; calculations from Westoff and Ochoa (1991).

Data Requirements

See the data requirements for the demand for limiting and demand for spacing indicators.

Data Source(s)

Population-based surveys.

Purposes and Issues

Total demand for family planning consists of the sum of demand for limiting plus demand for spacing, and thus provides a basis for estimating the total number of clients that would have to be served if all women currently desiring to either avoid further pregnancies or space their next birth were to seek family planning services. The reader is referred to the discussions of the "demand for limiting" and "demand for spacing" indicators presented earlier in this chapter for computational details and caveats.

A recent comparative study for 25 countries conducting a DHS during the 1985-1989 period illustrates the use of the total demand indicator and its partition into demand for limiting and spacing for program administration and evaluation purposes (Westoff and Ochoa, 1991). For the 25 countries as a whole, average total demand among women currently married or in union was estimated at approximately 63 percent. Total demand was considerable higher in Latin America and Asia (73 percent in each region) and North Africa (66 percent) than in sub-Saharan Africa (47 percent). Estimates of total demand ranged from a high of 81 percent in Brazil and Colombia to a low of just under 28 percent in Mali.

Large differentials were also observed with respect to demand for limiting versus spacing. In Latin America and Asia, approximately two-thirds of total demand was for the purpose of limiting, whereas in sub-Saharan Africa 50-60 percent of total demand was for spacing.

Estimates of total demand also provide key intermediate computations in the calculation of indicators of family planning program performance in satisfying demand for family planning services. Two such indicators are presented next in this chapter.

Indicator

UNMET NEED FOR FAMILY PLANNING

Definition

The number or proportion of women currently married or in union who are fecund and who desire to either terminate or postpone childbearing, but who are not currently using a contraceptive method.

The total number of women with an unmet need for family planning consists of two groups of women: (a) those with an unmet need for limiting, and (b) those with an unmet need for spacing.

Women with an unmet need for limiting are those who desire no additional children and are not currently using a contraceptive method.

Women with an unmet need for spacing are those who desire to postpone their next birth by a specified length of time (for example, for at least two years from the date of a survey) and are not currently using a contraceptive method.

The indicator is calculated as follows:

$$U = U_L + U_S$$

Where:

U = the number or proportion of women with unmet need for family planning,

U_L = the number or proportion of women with an unmet need for limiting,

U_S = the number or proportion of women with an unmet need for spacing.

Illustrative Computation

Estimate of unmet need for family planning, Ghana, 1988 (expressed as a proportion of women currently married or in union).

$$U = U_L + U_S$$

$$= 9.0 + 26.2$$

$$= 35.2$$

Source of data: Ghana Demographic and Health Survey, 1988; calculations from Westoff and Ochoa (1991).

Data Requirements

Responses to survey questions on:

- desire for additional children and, if so, the desired length of birth interval;
- current contraceptive use status;
- current fecundity, pregnancy, and amenorrhea status for women not currently using a contraceptive method;
- the wanted status (with respect to number and/or timing) of the current/last pregnancy for women currently pregnant or amenorrheic; and
- whether a contraceptive method was being used at the time of the current/last pregnancy.

Note: the use of the information in the final two items in the computation of the indicator is explained below.

Data Source(s)

Population-based surveys.

Purposes and Issues

This indicator provides information on the size of an extremely important population sub-group for family planning program management: women at risk of pregnancy with an apparent need for family planning services based upon their expressed desire to limit or space future births, but who are not using contraception. Such women are said to have an "unmet demand" or "unmet need" for family planning and are the logical primary target of program marketing/recruitment efforts.

The indicator may also be interpreted as the number of additional clients who would be using contraception (over and above the number of current users) if all women at risk of pregnancy and desiring to either terminate or postpone childbearing were to adopt contraception.

The indicator follows from the decomposition of total demand for family planning services into two components: "met demand" and "unmet demand" (or "unmet need"). Met demand includes women with demand for family planning (see the "total demand for family planning"

indicator immediately prior to this chapter for definitional and computational details) who are using a contraceptive method in order to achieve their reproductive goals; unmet need, or unmet demand, consists of women with an apparent demand for family planning who are not using contraception.

Following the procedure proposed by Westoff and Ochoa (1991), women are considered to be at risk of pregnancy in the present indicator if they are:

- of reproductive age and currently married or in union;
- fecund;
- not using a contraceptive method; and
- not currently pregnant or amenorrheic.

However, the following categories of women are not considered to have an unmet need for family planning, and they are thus excluded from the computation of the indicator:

- currently pregnant or amenorrheic women who were using contraception at the time they became pregnant with the current/last birth (these women are viewed as not in need since prior need was met through contraceptive use, although there would appear to be a need for a more effective method);
- currently pregnant or amenorrheic women whose pregnancy was reported as intentional; and
- fecund women who want their next child within the next two years.

The reader is referred to Appendix E for an illustration of the computation of the indicator.

Bongaarts (1991b) has proposed two modifications to the measurement procedure described above: (1) an adjustment to account for the fact that the satisfaction of need for spacing through contraceptive use will result in a reduction in the need for limiting to the extent that it postpones the date at which women reach their desired family size; and (2) an adjustment for what is perceived to be an overestimate of the need for spacing in the procedure described above. Bongaarts proposes using the estimates produced by the procedure described above as a starting point and introducing these two adjustments to compensate for the perceived problems. A comparison of the estimates from the two methods suggest that the Westoff procedure tends to produce estimates of the level of unmet need that are higher than those of Bongaarts by on average of about 5 percent (Bongaarts, 1991b; Westoff and Ochoa, 1991). The reader is referred to these references for further details on the two methods of computing the indicator.

Another possible refinement would be, as discussed earlier, to include all women of reproductive age reporting sexual activity in survey interviews in the computation of the indicator in populations where a significant share of childbearing occurs outside of recognized marriages/unions. One difficulty in doing this, however, lies in the identification of women not currently married or in union who are "at risk" of pregnancy. This would require the use of data on sexual activity among women not currently married or in union (Blanc and Rutenberg, 1990). Further research is needed on the accuracy of such data, given the sensitivity of survey inquiries in this area in some cultures.

Recently, some researchers have argued that the definition of "unmet need" should be broadened to include women using: (1) traditional contraceptive methods (on the grounds of high failure rates for such methods); (2) a theoretically effective method incorrectly or sporadically; and (3) a method that is unsafe or unsuitable for them (Foreit, 1992; Dixon-Mueller and Germain, 1992). The adoption of these alternative definitions would raise significantly the estimated numbers of women with unmet need for family planning in many developing country settings.

Indicator

SATISFACTION OF DEMAND FOR FAMILY PLANNING

Definition

The proportion of total demand for family planning at a given point in time that is being satisfied by current contraceptive use.

The indicator is calculated as follows:

PDS = C / D

Where:

PDS = the proportion of total demand for family planning that is being satisfied by current contraceptive use,

C = the number of women currently using a contraceptive method for the purpose of either limiting or spacing, and

D = the total number of women with a demand for family planning.

Note: separate estimates of the proportion of demand for limiting and spacing satisfied by current contraceptive use may be calculated by substituting estimates of demand for limiting and spacing and contraceptive prevalence by purpose (i.e., for limiting or spacing) into the equation above.

Illustrative Computation

Estimates of the proportion of demand satisfied by contraceptive use, Ghana, 1988 (for women currently married or in union).

Total Demand	For Limiting	For Spacing
$PDS = C / D$	$PDS_L = C_L / D_L$	$PDSs = C_S / D_S$
= 12.9/48.1	= 4.9/13.9	=8.0/34.2
= 26.8	= 35.3	= 23.4

Source of data: Ghana Demographic and Health Survey, 1988; calculations from Westoff and Ochoa (1991).

Data Requirements

Survey estimates of:

- the number or proportion of women of reproductive age currently married or in union who are currently using a contraceptive method (i.e., "met demand"), classified by whether women currently using a contraceptive method desire to either terminate childbearing or postpone their next birth if separate estimates of the proportion of demand for limiting and spacing being satisfied by contraceptive use are desired; and
- total demand for family planning, again broken down into demand for limiting and demand for spacing if separate estimates of the proportion of demand for limiting and spacing, respectively, being satisfied are desired (see the discussion of the "total demand for family planning" indicator in this chapter for computational details).

Data Source(s)

Population-based surveys.

Purposes and Issues

This indicator, based upon the decomposition of total demand for family planning into "met" and "unmet" components as described in connection with the previous indicator, provides an alternative to the contraceptive prevalence rate (CPR) as a population-based measure of the extent of contraceptive use. The present indicator is a somewhat more refined measure, however, in that it relates current contraceptive use, or "met demand," to total demand for family planning; that is, to the population of women who are at risk of pregnancy and have indicated a desire to restrict future fertility. The indicator thus provides a more specific measure of program success in satisfying existing demand for services than does the CPR.

The reader is referred to the observations on the rationale for broadening the universe of the measure to include women who are not currently married or in union and/or those using traditional or inappropriate contraceptive methods made in connection with the indicators presented earlier in this chapter.

- Number of visits to service delivery point(s)
- Number of acceptors new to modern contraception
- Number of acceptors new to the institution
- Number of new segment acceptors
- Couple–years of protection (CYP)
- Method mix
- User characteristics
- Continuation rates

INDICATORS TO MEASURE SERVICE UTILIZATION

Most administrators are interested in evaluation to measure the results obtained from their program. A program manager might want data on the utilization of specific services (i.e., one or more service delivery points), whereas a director of a national family planning program would be more interested in a nationwide coverage (i.e., the percentage of women at risk of pregnancy currently using contraception).

These two topics, service utilization and contraceptive use, are closely linked, in that most clients will come to use a family planning method through one of the family planning delivery mechanisms: a clinic (or clinic–based program, e.g., postpartum program), a community–based distribution program, a commercial/social marketing program, or a private source. However, not all clients who utilize a given facility will actually use the method that they take home from their visit; and not all who initiate the method will stay with it (Elkins and Nordberg, 1977; Morris and Anderson, 1982). Thus, the distinction between service utilization and contraceptive use remains a valid one.

Service utilization tends to be evaluated using program–based measures, since programs collect data on the number and types of clients serviced, methods sold or distributed free of charge, characteristics of the clients, and so forth (Chandrasekaran, 1975). While the number of clients served can be used to arrive at a rough estimate of coverage within the populations, such estimates may be flawed if one does not have accurate statistics on the size of the target population, if non–residents of the catchment area make use of the services, or if multiple visits by a single client are not discounted (Morris and Anderson, 1982; ICDDR–B, 1989).

Because of these problems in estimating coverage from service statistics, contraceptive prevalence is most accurately reported from population–based data. To achieve consistency between the blocks of the conceptual framework (see Figure I–1 in Chapter I) and the titles of the chapters in the Handbook, program–based indicators are presented in this chapter (VII), whereas the population–based indicators for measuring contraceptive practice are presented in the next chapter (VIII).

This conceptual distinction is blurred by the fact that most of the indicators of service utilization and contraceptive use presented in the Handbook can be obtained from either service statistics or from surveys. For example, it is possible for a clinic to calculate continuation rates among its client population based on data of clients returning to that clinic accompanied by a follow–up of dropouts. Continuation rates of other types (e.g., use of any method from any source) can be measured from population–based data (e.g., using the calendar of the DHS).

To the extent that a given indicator is commonly calculated from both types of data, it is listed and cross–referenced as a measure of service utilization (Chapter VII) and of contraceptive practice (Chapter VIII). By contrast, in those cases where the indicator can theoretically be obtained from either type of data but rarely is, it is listed according to the source of data more commonly used.

Readers should also be alerted to the fact that depending on the source of data, service utilization may be classified as an output (program–based) or an effect (population–based).

Readers will note that the Handbook includes three variations of the indicator “new acceptor.” This indicator is among the most inconsistently defined at the field level, in part because different definitions serve different (and valid) purposes. Thus, rather than force a standard definition of “new acceptor” which will invariably be at odds with current usage in some countries or programs,

we have attempted instead to provide clear definitions for three variations on "new acceptor" and to indicate the utility of each one.

Two indicators conspicuously missing from this chapter on service utilization are the number of continuing users and the number of dropouts. While these measures may serve some purpose in heightening staff awareness regarding client continuation, they are viewed as less useful than the continuation rate among clients for evaluation purposes.

Moreover, there are several problems in measuring dropouts. First, the definition of dropouts has not been consistent among programs (varying primarily in terms of the treatment of those who return after interrupted use and allowances for "grace periods" of differing lengths). Second, the level of effort required to properly monitor dropouts often exceeds staff capacity, resulting in poor data quality on this indicator. Given that the number of dropouts is required to calculate the number of continuing users, the data problems regarding dropouts also affect the accuracy of figures on number of continuing users. If information on the number of dropouts is needed, it can be obtained from continuation rate computations.

The authors of this Handbook concur with members of an earlier USAID Task Force on Improving Family Planning Program Performance Indicators that number of dropouts need not be retained in the battery of key indicators for service utilization (USAID Task Force on Standardization of Family Planning Program Performance Indicators, 1987). This is not to say, however, that service providers should not follow–up on dropouts as a means of improving service delivery; rather, it is recommended that programs not rely on the number of dropouts as a key program evaluation indicator in view of the data collection problems noted above.

Indicator

Number of Visits to Service Delivery Point(s)

Definition

The total number of visits made by clients to a service delivery point (or to all SDPs within a system) in a given reference period (e.g., one year), preferably by purpose (contraceptive supply, counseling, referral, other).

Data Requirements

Enumeration of the number of client contacts at an SDP, preferably by type of visit, aggregated over SDPs to get a total for the system.

Data Source(s)

Service statistics collected at program facilities.

Purpose and Issues

This indicator measures the volume of activity undertaken at SDPs. It is a substitute for measures of numbers of acceptors or users, where such data are not available or are considered unreliable. This information can be combined with other data to obtain crude measures of productivity (e.g., the mean number of client visits per full–time service provider or per clinic). Its main advantage is its simplicity; it is easy to collect and uses simple definitions. The breakdown of visits by type of visit provides information on the nature of the provider–client transaction.

There are several limitations to this indicator. First, this measure does not provide information on the number of clients, since a single individual can make multiple visits during a given period. Second, as generally applied in field settings, it doesn't distinguish new versus continuing users; thus, it doesn't indicate the success of recruiting new acceptors as opposed to increasing continuity among current users.

Indicator

Number of Acceptors New to Modern Contraception

Definition

The number of persons who accept for the first time in their lives any (program) method of contraception; to be reported for a defined reference period (e.g., one year).

Data Requirements

Counts of persons accepting any (program) method for the first time in their lives during a one year period.

Data Source(s)

Service statistics; surveys (possible but not common).

Purpose and Issues

This indicator measures the ability of the program to attract new clients to its services from an untapped segment of the population. The measure eliminates the problem of counting as "new" those clients who switch from one source to another for reasons of convenience or cost. As an indicator, it may also reflect the success of special communication programs or other interventions (e.g., social marketing projects) aimed at increasing service utilization among those previously not reached by the program. However, in this latter case, one must be mindful that some of the new acceptors might have obtained the same or another method from an alternate source (e.g., the un–subsidized pharmacy sector) if the special intervention had not taken place.

"Program method" refers to methods made available through established family planning programs: pill, IUD, the NORPLANT® implant, injection, condom, spermicides, diaphragm, tubal ligation, vasectomy, and lactational amenorrhea method (LAM), if used under program supervision. Thus, a young woman who formerly obtained condoms from the pharmacy would not be a new acceptor. By contrast, a client who to date has depended on withdrawal would be a new acceptor, since this is not a program method.

The "number of acceptors new to modern contraception," defined as first–time use in the life of the individual, reduces the ambiguity associated with the more general term "new acceptor" and avoids a duplication of cases that may result when substitution occurs.

This indicator can be obtained from survey data as well (e.g., from the "calendar" used in the DHS or other data collection tools for obtaining contraceptive histories retrospectively). However, in the context of surveys, total current use rather than "new use" is likely to be of greater interest to most users of the information.

Indicator

Number of Acceptors New to the Institution

Definition

The number of persons during a defined reference period (e.g., one year) who accept a contraceptive method from a particular institution for the first time, although they may have previously used methods obtained elsewhere.

Data Requirements

Counts of persons during the reference period who accept a contraceptive method from a specific institution.

Data Source(s)

Service statistics (preferred); surveys (possible but not common).

Purpose and Issues

This indicator measures the ability of the institution to attract clients specifically to its services, even if they have used another source previously.

Most institutions are not evaluated on how family planning is doing in the country as a whole, but rather on how well they as an institution are succeeding in attracting new clients. Thus, even if clients are drawn from other sources to the program, it reflects well on their institutional performance. This indicator is also important in the context of financial sustainability, given that private sector providers may be partly evaluated on the basis of their success in attracting clients who were previously using public sector services.

This variation on "new acceptor" is widely used in that it has administrative implications. Clients who receive services for the first time at a particular institution may be required to fill in additional paperwork, have a client record established, pay an enrollment fee, and so forth. Thus, the administrative requirements of the program yield data that are also useful for program monitoring.

One source of inconsistency regarding this indicator involves the treatment of clients who discontinue use of family planning services in a given institution but return at a later date. For example, some organizations keep records for a fixed time period (e.g., five years). In this case, a person returning after a five–year lapse would be counted as a new acceptor. In other institutions, one can be "new to the institution" only once. We recommend the latter usage of the term, while recognizing that administrative constraints cause some organizations to adopt the former.

One common variation on the indicator that is not listed separately in this Handbook is "new to the SDP." This measure is in fact used in some programs, since it simplifies the operational definition of new acceptor; the receptionist or other staff member need only ask the client, "Have you ever been to *this* clinic (dispensary, health post, or other type of service delivery point)?" Thus, a client who moved from one neighborhood to another and changed government health centers in the process would be considered "new to the SDP" at the health center in the new neighborhood. Although this alternative indicator is used in some programs because it is simple to understand and apply, it invites a duplication of cases, especially in countries with a high level of residential mobility. For this reason, it is not presented as one of the recommended indicators in this Handbook.

Indicator

NUMBER OF NEW SEGMENT ACCEPTORS

Definition

The number of persons who are initiating a new segment of contraception during a defined reference period (e.g., one year); that is, they are "new" to a particular contraceptive method or to the source institution during the defined reference period. The individual may have used a method, including this one, before, and may have used this source before, but was not using this method and this source at the time of acceptance.

Data Requirements

Counts of persons accepting a new method from a particular source during the reference period.

Data Source(s)

Service statistics (preferred); surveys (possible but not common).

Purpose and Issues

This variation on the definition of "new acceptor" serves two important purposes. First, it gives program managers a means of assessing the service delivery requirements for their program in a given calendar year. In the case of sterilization, IUDs, and NORPLANT® implants, the services that providers need to give to clients are largely associated with the first visit when they adopt the method. Second, it is this variation on the definition of new acceptor that is used in the Target Cost Model (Stover et al., 1991).

Under this definition, an individual adhering to any of the following criteria would be classified as a new segment acceptor if he or she:

- has never used contraception before;
- has used the same method but at a different source (institution), in the past or immediately prior to the current acceptance;
- has used the same method from the same source in the past, but was not using it immediately prior to the current acceptance; or
- was using contraception from the same source prior to the current acceptance, but changed to a new method.

If the client has not previously used the method, then new segment acceptor is the equivalent to an acceptor new to the method. However, new segment acceptor also includes women who previously used the method, discontinued for a period (for example, for pregnancy), and then resumed use of the same method.

If a client within the national family planning program changes clinics but her records follow her, then this is not a change of source. However, if different clinics within a larger system have separate record–keeping systems, this case would constitute a change of source.

Indicator

COUPLE–YEARS OF PROTECTION (CYP)

Definition

The estimated protection provided by family planning services during a one year period, based upon the volume of all contraceptives sold or distributed free of charge to clients during that period. The CYP is calculated by multiplying the quantity of each method distributed to clients by a conversion factor, which yields an estimate of the duration of contraceptive protection provided per unit of that method (Wishik and Chen, 1973; Centers for Disease Control, 1985). The CYPs for each method are then summed over all methods to obtain a total CYP figure.

The conversion factors currently in use in the USAID–system are those endorsed by the A.I.D. Task Force on Standardization of Family Planning Program Performance Indicators (1991):

Method:	*CYP Per Unit*
Oral Contraceptives:	*15 cycles per CYP*
Cu "T" 380–A IUD:	*3.8 CYP per IUD inserted*
NORPLANT® implant:	*3.5 CYP per implant*
Condoms:	*150 units per CYP*
Vaginal foaming tablets (Conceptrol, Neo–Sampoon):	*150 tablets per CYP*
Sterilization (male or female):	*10 CYP per procedure*
Depoprovera (injectable):	*4 doses (ml) per CYP*
Noristerat (injectable):	*6 doses per CYP*
(monthly injectable):	*12 doses cyclofem per CYP*
Natural Family Planning (NFP):	*2 years per trained, confirmed adopter*
Lactational Amenorrhea Method (LAM):	*4 active users per CYP*[1]

[Note: an illustrative computation of this indicator is provided at the end of the discussion of the indicator.]

Data Requirements

Quantities of pills, condoms, and spermicides distributed to clients; numbers of IUDs and NORPLANT® implants inserted; number of injections administered; number of sterilization operations performed; number of trained, confirmed clients of NFP; number of LAM clients during the reference period.

Data Source(s)

Service statistics or logistics management information system.

Purpose and Issues

CYP measures the volume of program activity. It is used by program managers and donor agencies to monitor progress in the delivery of contraceptive services at the program and project levels. Because USAID and IPPF require the organizations they support to report CYP, this measure is currently one of the most widely–used indicators of output in international family planning programs.

There are several advantages of the indicator:

- it can be calculated from data that are usually collected on a routine basis through programs or projects, thus minimizing the data collection burden;
- these data can be obtained from all the different service delivery mechanisms (clinics, CBD, social/commercial marketing); and
- the CYP calculation is relatively simple to do.

The principal disadvantages of the indicator are that:

- it is not intuitively easy to understand by those outside the field;
- one cannot ascertain the number of individuals that are represented by CYP. For example, if a program administers 10,000 injections of

[1] LAM was not included on the list developed by the Task Force; this conversion factor has been supplied by Labbok, 1992.

DepoProvera, this amount is equivalent to 2,500 CYP. Theoretically, this represents 2,500 women protected for 12 months each; however, in fact it may be 5,000 women covered for 6 months each or 10,000 women covered for 3 months each; and

- the validity of the assumptions underlying the choice of conversion factors has been widely questioned (Stover et al., 1993).

Regarding the calculation of CYP for long–term methods, in most programs the entire amount is "credited" to the calendar year in which the method is accepted. For example, if a family planning program performed 100 VSC procedures in a given year, it would credit all 1000 CYP (100 procedures x 10 years/each) to that calendar year, even though the protection from those procedures would in fact be realized over that and the next nine years. An alternative approach is to "annualize" this projection, allocating it over a ten year period. The same principle applies to IUDs and the NORPLANT® implant. Although the first approach (of crediting the full amount of CYP in the calendar year of acceptance) has been harshly criticized, nonetheless it represents current practice in most programs that report CYP, probably because it is easier to apply.

Ideally, CYP should be based on the volume of contraceptives that are delivered to clients who will presumably use them, not on those delivered to facilities where they may remain unused in cartons or on shelves. However, in some projects such as social marketing, it may be impossible to monitor the exact numbers reaching the hands of clients. Rather, the only means of calculating CYP is to base it on the volume of contraceptives delivered to the retailers in question. Given that retailers are unlikely to stock products that do not move readily, it is probable that (after an initial shipment) most contraceptives sold to retailers will make their way into consumers' hands. However, in those instances where the calculation of CYP is based on the volume of products delivered to retailers, not directly to the clients or customers themselves, this detail should be made clear to the users of the information.

The conversion factors listed above are those proposed in January 1991 by the Task Force on Family Planning Program Performance Indicators, in part in response to the commentary "What's Wrong with CYP" (Shelton, 1991). There is still considerable debate over the appropriateness of these factors, since several of them were not derived on an empirical basis.[2] Under The EVALUATION Project, an attempt is underway to examine existing empirical evidence (on wastage, mean duration of use, frequency of sexual relations, consistency of use, and the non–contraceptive use of condoms), to assess whether in fact these conversion factors are appropriate and if not, what factors would be preferable (Stover et al., 1993). Relevant findings will be incorporated into the conversion factors presented in the next version of the Handbook.

Illustrative Computation

CYP, based upon conversion factors given in text.

Method	Quantity	CYP
Oral contraceptives	5,022	334.8
IUDs	87	330.6
Condoms	62,810	418.7
Vaginal tablets	3,900	26.0
Tubal ligations	13	130.0
Depoprovera	1,277	319.3
TOTAL		1,559.4

[2] Some have argued that in light of the controversy surrounding the conversion factors for calculating CYP, one might better monitor the volume of each contraceptive sold. This approach would be appropriate to a program that relies on only one or two methods, but it is less practical for the more conventional program that offers six to eight different methods and thus would have to monitor each separately.

Indicator

METHOD MIX

Definition

The percentage distribution of contraceptive users (or alternatively, of acceptors) by method.

Data Requirements

Number of users (of acceptors) by method.

Data Source(s):

Service statistics; (for a population–based measure, see Chapter VIII).

Purpose and Issues

The method mix provides a profile of the relative level of use of different contraceptive methods. A broad method mix suggests that the population has access to a range of different contraceptive methods. Conversely, method mix can signal: (1) provider bias in the system, if one method is strongly favored to the exclusion of others; (2) user preferences; or (3) both.

Because of the problems of monitoring the number of current users based on service statistics, method mix is generally based on acceptors, not current users, when measured at the program level. The two would not yield the same distribution, since user data would reflect the accumulation of long–acting methods from previous years.

Similarly, one would expect some discrepancy on method mix calculated from program statistics versus surveys, even in programs with reliable data. (The reason is that program–based statistics reflect activity in the calendar year under study, whereas the survey results include continuing users of long–acting methods who adopted them in previous years and have not needed or chosen to return to the clinic in the calendar year under study). In addition, survey data include folk methods, non–program methods (e.g., withdrawal), and program methods that are also available from non–program sources (e.g., pills from pharmacies).

Despite considerable discussion of method mix, there has been relatively little in the published literature on what constitutes a desirable method mix (an exception being Hutchings et al., 1987). It is generally felt that a program should respond to the changing needs of the population at different stages in the reproductive life cycle, offering reversible methods for those who desire to space and permanent methods for those who have completed their desired family size. Thus, programs that offer no permanent methods or that overemphasize permanent methods are subject to criticism. Yet within the category of reversible methods, the distribution of acceptors by type of contraceptive will vary by availability of specific methods, costs, local preferences, and other factors, making it difficult to generalize regarding desirable method mix.

Indicator

USER CHARACTERISTICS

Definition

A socio–demographic profile of current users of contraceptive methods relevant to program planning and/or marketing.

Relevant characteristics include: age, parity, urban–rural residence, economic status, ethnicity, and other factors judged important in the context of a specific country.

Data Requirement

Data on age, parity, and other characteristics of users (or if not available, of acceptors).

Data Source(s)

Service statistics; population–based surveys (see Chapter VIII).

Purpose and Issues

This indicator measures the characteristics of clients receiving contraceptive services at service delivery points in the program. In those cases where the program is designed to reach subgroups with specific socio–demographic characteristics (e.g., lower parity women), it indicates the extent to which the program is reaching its target population. Similarly, large urban/rural differentials in a public program may signify that it is not reaching key target populations.

Ideally, one would like to have this type of information on all current users. However, due to the difficulty of monitoring current users, an alternative approach is to obtain data on acceptors as they enter the program.

Program statistics on user characteristics allow a program to monitor how its client population changes over time. Typically, a new program in a low prevalence country will initially attract older, high–parity women whose felt need for the service is great. Over time, the mean age and parity of the clients decreases, as family planning becomes more socially acceptable and attracts couples earlier in their reproductive lifetime.

Indicator

CONTINUATION RATES

Definition

The probability that an acceptor of a contraceptive method will still be using some contraceptive method after a specified period of time (e.g., one year). This is known as the "all–method" continuation rate (United Nations, 1991; Jejeebhoy, 1989).

Data Requirements

The number of acceptors who initiate contraceptive use at a given point (or during a given period of time) and the length of time that each individual continues to use the method (or a substitute method). Based on this information, one can calculate the percentage who have continuously used for a specific duration (e.g., 12 months, 18 months, etc.), as well as the mean or median duration of use.

Data Source(s)

Program–based: client records accompanied by a follow–up study of program drop–outs.

Population–based: surveys with retrospective contraceptive use histories or calendars (see Chapter VIII).

Purpose and Issues

Contraceptive continuation rates provide a useful summary measure of the overall effectiveness of program services in enabling clients to sustain contraceptive use. The indicator has also been proposed as a measure of the effectiveness of service follow–up mechanisms.

Despite the current interest in contraceptive continuation in relation to quality of care, accurate data on continuation tend to be difficult and expensive to collect through programs.

Continuation rates may be estimated by interviewing "real" cohorts of acceptors (clients who accept a contraceptive method in the same month) at one or more fixed intervals of time after acceptance (e.g., 12 months) in order to determine the percentage still using the method after that elapsed time. This provides a continuation rate that is intuitively simpler to calculate than the "life table" rates that are calculated from survey data (see Chapter VIII for further discussion of the use of survey data to calculate continuation rates).

However, continuation rates derived from facility–based sources of data tend to be characterized by: (a) selection bias, since the characteristics of clients attending clinics may be different from contraceptive users in the general population; and (b) significant numbers of acceptors who are "lost–to–follow–up." As a result of the latter problem, estimates based on those individuals with whom the program has maintained contact may not accurately measure continuation rates even among the population of program users. Lost–to–follow–up problems are difficult to eliminate completely, and undertaking the periodic follow–up surveys required to minimize the resulting bias can be costly.

An additional problem with estimates of continuation rates (and of failure rates, described in Chapter VIII) based upon clinic/facility data is that such estimates do not capture the experience of contraceptive users who obtain supplies from non–clinic sources (e.g., from commercial outlets). To the extent that clinic–based service providers are able to provide useful counseling on the proper use of methods accepted, it is likely the level of proper use of contraceptive methods among clients of clinic–based services will be different from that of commercial sector clients.

Two alternative continuation rates are sometimes used:

- first–segment continuation rate: the probability that an acceptor of a particular method will continue to use that method without interruption for a specified period of time; and
- first–method continuation rate: the probability that an acceptor of a particular method will continue to use that method for a specified period of time without changing methods or becoming pregnant, allowing for short intervals of discontinuation (Jejeebhoy, 1989).

The current trend among service providers concerned with quality of care is to encourage method switching if the first method proves unsatisfactory, on the premise that this practice will result in greater client satisfaction and increased continuation rates. In light of this, the "all–method continuation rate" would seem more appropriate than the two alternative "first method" rates presented directly above in assessing continuation in a program.

For further discussion see Chapter VIII.

Chapter VIII

Contraceptive Practice

- Contraceptive prevalence rate (CPR)
- Number of current users
- Level of ever (past) use
- Source of supply (by method)
- Method mix
- User characteristics
- Continuation rates
- Use failure rates

INDICATORS TO MEASURE CONTRACEPTIVE PRACTICE

As mentioned in the previous chapter, service utilization and contraceptive practice are closely linked yet conceptually distinct. In this chapter we present measures of contraceptive practice that use population–based data.

The World Fertility Survey (WFS) was the first of the large–scale surveys designed to obtain data on a nation–wide representative sample of women of reproductive age regarding fertility and family planning (Kendall, 1979). Although they provided a wealth of information on the determinants of fertility, they did not provide sufficient data for the design and evaluation of family planning programs; moreover, they required considerable time to process and analyze the data, detracting from the utility of the information to program administrators.

An alternative survey instrument was developed for programmatic purposes: the Contraceptive Prevalence Survey (CPS). First implemented in El Salvador in 1975, it proved a useful source of data on the family planning program (Morris et al., 1981). The initial series of CPS surveys conducted by the Centers for Disease Control was then expanded into a larger CPS project, awarded by USAID to Westinghouse Applied Health Systems (now IRD Macro) in the late 1970s. At the same time the CDC has continued to conduct surveys on family planning (FPS) and reproductive health.

In the 1980s the CPS gave way to an expanded format, the Demographic and Health Surveys (DHS), which included: (a) the programmatic family planning data from the CPS, (b) the detailed demographic data from the WFS, and (c) a module on the health status of women and their children under five. Over 40 countries have participated in either the DHS or FPS. A total of 31 countries have now conducted two or more surveys, including WFS, CPS, DHS, FPS, and other comparable surveys (London et al., 1985; Robey et al., 1992). (For simplicity, the term "DHS" is used throughout the Handbook, but should be interpreted as "DHS–type surveys" including those conducted by CDC and other national–scale demographic or reproductive health surveys conducted under other auspices.)

These surveys provide the population–based data mentioned so frequently throughout this Handbook. Numerous surveys of a similar nature have been carried out for selected areas (regions or cities) in connection with operations research projects or for other research/evaluation purposes, which also provide population–based estimates of contraceptive prevalence.

Prior to these large–scale surveys, demographers generally attempted to evaluate the coverage of family planning programs by estimating the numerator from service statistics, the denominator from an estimate of the catchment area based on census or other demographic data. However, surveys have come to be the method of choice for estimating coverage of programs (e.g., contraceptive prevalence). Surveys eliminate the problem of double–counting (if a woman has visited more than one facility to obtain contraception); they include individuals who obtain their contraceptives from other than program sources (e.g., a private physician); and they identify individuals who use methods not requiring program contact (e.g., withdrawal) (London et. al, 1985). One limitation of surveys is possible undercounting of use (if women choose to answer "no" to a yes/no question on current contraceptive use, either to get on with the survey or to avoid detection if practicing in secret).

The following indicators summarize the most useful pieces of information available from large–scale surveys for measuring contraceptive practice.

Indicator

CONTRACEPTIVE PREVALENCE RATE (CPR)

Definition

The proportion of women of reproductive age who are using (or whose partner is using) a contraceptive method at a particular point in time, often reported for women married or in sexual union.

The indicator is calculated as follows:

CPR = U / P

Where:

U = the number of women using a contraceptive method at a given point in time, and

P = the number of women of reproductive age (or alternately, women of reproductive age currently married or in union).

Illustrative Computation

Contraceptive prevalence rate among women aged 15–49 years of age, Kenya, 1989

All women:	Women currently married or in union:
CPR = 1,659 / 7,150	CPR = 1,282 / 4,765
= .232	= .269

Source of data: Kenya Demographic and Health Survey, 1989.

Data Requirements

The total number of women of reproductive age, by marital status (optional); and of these, the number that are currently using a contraceptive method.

Data Source(s)

Population–based surveys.

Purpose and Issues

The CPR provides a measure of population coverage of contraceptive use, taking into account all sources of supply and all contraceptive methods; it is the most widely reported measure of outcome for family planning programs at the population level.

Technically speaking, CPR is a ratio, not a rate. Prevalence is measured by a ratio and incidence by a rate. For a given year, contraceptive prevalence measures the proportion of women of childbearing age in union who use a form of contraception. To obtain a true contraceptive use rate, the measure should have in the denominator the population at risk (of pregnancy), i.e., sexually active women who are not infecund, pregnant, or amenhorreic. The numerator would have the number of contraceptive users from that population. It should be stressed that this point is included for informational purposes only. The international population community uses the term "contraceptive prevalence rate" as defined above, and thus this Handbook endorses this practice to assure consistency.

The convention in reporting contraceptive prevalence is to base this calculation on women married or in sexual union (even though most DHS–type surveys ask questions of contraceptive use to women of reproductive age, regardless of their marital status). In countries with relatively little sexual activity outside marriage for women, basing prevalence estimates on women in sexual union captures the population at risk of pregnancy. However, in countries with the widespread practice of sexual activity outside of stable sexual unions, a prevalence estimate based on women in union only would ignore a considerable proportion of current users. Thus, researchers are continuing to study the appropriateness of basing contraceptive prevalence on all women versus those in stable unions.

Whereas the CPR may theoretically be derived from service statistics on numbers of current users and estimates of the population at risk, current practice is to rely upon population–based sample surveys in order to minimize the problems associated with maintaining a running count of current users and obtaining accurate population estimates. (The problems include incomplete data, double–counting of users who enter the service delivery system at more than one point, purposeful inflation of service statistics, and poor quality of data due to other activities competing for the

attention of those recording the information, to name the primary ones.)

The DHS is currently the main source for obtaining national level estimates of prevalence. As mentioned above, "DHS" is used in this Handbook to mean "DHS–type surveys": the actual DHS, the Family Planning Surveys conducted by CDC, and other large scale national surveys conducted by the countries themselves under other auspicies (e.g., in Algeria, Bangladesh, China, Hong Kong, India, Singapore, South Africa, South Korea, Taiwan, Turkey, and Vietnam). Smaller scale and/or more focused surveys may also be used to estimate the CPR, the essential ingredient for obtaining scientifically–sound estimates being the use of probability sampling methods. Estimates may also be obtained by adding relevant questions to surveys on other topics (e.g., health program prevalence or coverage surveys), given appropriate sampling methods and sample sizes.

Indicator

NUMBER OF CURRENT USERS

Definition

The number of women (or their partners) of reproductive age who are estimated to be using a contraceptive method at a given point in time; this can be reported by type of method, region, type of source, or other relevant variable.

Data Requirements

Counts of women (or their partners) using a contraceptive method at a particular point in time.

Data Source(s)

Population–based surveys (preferable); service statistics (see Chapter VII).

Purpose and Issues

Number of current users provides a summary measure of total program service volume. Prior to the introduction of population–based surveys to evaluate family planning programs, the number of current users was calculated from program–based statistics, and it constituted a widely used indicator of program output. It is appealing in that: (a) it is intuitively clear, (b) it directly measures what programs are trying to generate: users of contraception, and (c) it can be used to compare the performance of a given program over time or across regions.

Although conceptually clear and seemingly easy to collect, counting the number of current users from program records has proven to be a labor–intensive, time–consuming activity, especially in large programs that do not have a computerized client record system (including most CBD efforts). Specifically, it requires: (1) estimating the period of time for which each individual client should be considered active, based on the amount of contraceptive supply he/she receives, and (2) summing the total number of persons estimated to be current at a given point. For this reason many programs have discontinued the monitoring of current users and instead report couple–years of protection (see previous chapter) as a measure of program activity.

One can estimate the number of current users in the entire population (not just program users) by multiplying the estimated number of women of reproductive age in sexual union (based on census data or U.N. estimates/projections) by the contraceptive prevalence rate (from a survey). This figure can then be partitioned by source of supply available from the survey data (although many respondents may not be able to give accurate and complete information on the source). This estimate of current number of users is useful to program managers for anticipating the service needs of their program in terms of providers and supplies in different regions of the country. Nonetheless, most researchers and evaluators are more interested in contraceptive prevalence (a percentage of all married women) than in an estimate of the absolute numbers using a method.

Indicator

LEVEL OF EVER (PAST) USE

Definition

The proportion of women of reproductive age who have ever used a contraceptive method, including those currently using one.

Data Requirements

The number of women of reproductive age who report having ever used a contraceptive method (including those currently using one), the total number of respondents, and marital status (optional).

Data Source(s)

Population–based surveys.

Purpose and Issues

Ever use of contraception provides a crude measure of the extent to which a given population has experimented with methods of contraception, i.e., that have first–hand knowledge of contraception by having tried it at some point. When a new family planning program is introduced into a country where contraceptive prevalence is very low, this indicator can be useful in assessing the extent to which the program has "reached" (caused some response on the part of) this population.

The DHS asks the question, "Have you ever used anything or tried in any way to delay or avoid getting pregnant?" of all women reporting having heard of at least one method to delay or avoid pregnancy. The results may be reported for all women or for women that have ever been in a sexual union. In countries with considerable contraceptive use for premarital sexual activity, the choice of reference group may substantially affect the results.

Data on past use of contraception is generally collected as part of the history taken for new users of contraceptive services. However, this information tends to be used in better understanding the client's needs and experiences to date. It is rarely reported for evaluation purposes, and thus program statistics are not mentioned as a source of data above.

Comparisons of ever use and current use provide potentially useful information on contraceptive continuation. For example, if 80 percent of the population has ever tried a modern method, yet only 10 percent currently use a method, this would suggest a high degree of interest in contraception but seeming dissatisfaction with the program services or methods available. Such data would signal the need to further explore this situation. However, the ratio of ever use to current use is affected by numerous factors, rendering it of little value for evaluation purposes, except in relatively extreme situations to signal a potential problem, as noted above.

Indicator

SOURCE OF SUPPLY (BY METHOD)

Definition

The percentage distribution of the types of service delivery points cited by users as the source of their contraceptive method (if more than one source, then the most recent one).

Data Requirements

Number of respondents currently using contraception, the type of method used, and the source of supply of their method (most recently).

Data Source(s)

Population–based surveys.

Purpose and Issues

This indicator is useful to family planning program officials to show where contraceptive users obtain their supplies, for both evaluating program effectiveness and forecasting procurement needs. It is particularly appropriate to countries trying to shift the burden for family planning services from the public to the private sector. For example, the DHS–type surveys yield information on the percentage of modern method prevalence accounted for by the private sector.

In most countries the source of supply will vary substantially by type of method. VSC, IUDs, and the NORPLANT® implant require a clinic–based facility (including mobile clinics). Pills are available through clinics, in addition to commercial and CBD outlets. DepoProvera, once a clinic–based method, is being introduced into CBD programs on an experimental basis. Condoms and spermicides can be dispensed from any type of facility. Thus, data on source of supply are particularly useful when classified by method.

"Source of supply" yields two types of information: type of facility and type of sector (public/private). Type of facility generally includes hospital, health center, family planning clinic, mobile clinic, pharmacy, field worker, private doctor, and shop, among others. Sector distinguishes between governmental programs and those in the private sector (including the local family planning association, commercial retailers, private physicians, and other private providers). Ideally, data on source of supply should yield the proportion of contraceptive use attributable to the government program, the private family planning association, the private sector (pharmacies, private doctors), and other relevant sources.

However, the distinction between public and private is often difficult to make, especially in countries with multiple sources of contraception. The respondent may incorrectly identify a given clinic as a government clinic, when in fact it is private (or she simply may not know). A private physician may in fact be participating in a subsidized program to offer low cost services to specific target groups. In response to this problem, the DHS III questionnaire will provide a line for entering the actual name of the facility. Subsequent to the interview, a member of the research team will code the place mentioned according to the correct classification, based on master lists of SDPs. Those SDPs not on the list can be classified at a later date by contacting key informants from the area.

Indicator

METHOD MIX

Definition

The percentage distribution of contraceptive users by method.

Data Requirements

Number of current users by method.

Data Source(s)

Population–based surveys; service statistics (see Chapter VII).

Purpose and Issues

As mentioned in the previous chapter, method mix indicates the distribution of contraceptive use across different methods of contraception. It can reflect provider bias, supply problems, and client preferences.

Data on method mix, obtained both from surveys and from service statistics, are essential in the forecasting of commodity and service needs in the future. These data provide estimates of the current number of clients using these methods; changes in method mix over time are useful in projecting future trends (such as steadily increasing use of female sterilization in a number of countries).

In the case of method mix, it is not a question of which source of data is better: program– versus population–based. Both are used in forecasting the future contraceptive needs of a country. Survey data would be considered a more reliable means of assessing preferences for specific methods, given that they include clients from both the public and private sector, in addition to those using a non–program method such as withdrawal. However, one must be mindful that survey data (e.g., the DHS) may have a large standard error, especially where the percentage using a specific method is very low.

Indeed, survey data and service statistics do not always agree, a situation that can arise from inflated service statistics, wastage in the system, or the sale of products outside the target area for the program (e.g., across borders). To reconcile such discrepancies, special studies may be conducted.

Indicator

USER CHARACTERISTICS

Definition

A socio–demographic profile of current users of contraceptive methods relevant to program planning and/or marketing.

Relevant characteristics include: age, parity, urban–rural residence, economic status, ethnicity, and other factors judged important in the context of a specific country.

Data Requirement

Data on age, parity, and other characteristics of users (and in surveys, of non–users, for purposes of comparisons).

Data Source(s)

Population–based surveys; service statistics (see Chapter VII).

Purpose and Issues

This measure indicates the type of persons using contraception in a given country (or sub–region). User profiles based upon service statistics reflect the characteristics of clients receiving contraceptive services at program service delivery points, whereas profiles based upon survey data reflect the characteristics of users obtaining contraceptive services and supplies from all service and distribution sources. Thus, a comparison of the characteristics of the user population to those of the general population provides insight into the types of people being reached (e.g., better educated, more likely to have paid employment, higher parity, etc.).

From survey data, it is also possible to compare the characteristics of users versus non–users, and thus to identify sub–populations not being reached by the existing program. Surveys have the added advantage over service statistics that more detailed information on the characteristics of contraceptive users is normally available.

This indicator asks the question: "Of all individuals using contraception, what is the breakdown by _____ (urban/rural residence, educational level, religion, ethnic group, etc.)?" For example, one might find that of all contraceptive users in a given country, 70% were Catholic and 30% Evangelical. If in that country, 70% of the population were Catholic, 30% Evangelical, then this breakdown would simply mirror the composition of the general population (see Panel B in the example below).

The indicator on user characteristics should not be confused with the question asked in many research studies: "How does contraceptive use differ by subgroup in the population?" (which compares the percentage using contraception for the different categories of the explanatory variable, e.g., urban/rural residence, educational level, religion, ethnic group, etc.). To continue with the above example on religion, one might find that 40% of the married women of reproductive age in the total population used contraception, and that there was no difference by subgroup (see Panel C, below).

Examples:

Panel A: Absolute Numbers: Cross tabulation of modern contraceptive use by religion

Uses Modern Contraception:	Religion		
	Catholic	Evangelical	Total
Yes	280	120	400
No	420	180	600
Total	700	300	1000

Panel B: Among users, what is the breakdown by religion? (What religion do users tend to be?)

(Answer: 70% Catholic, 30% Evangelical)

Uses Modern Contraception:	Catholic	Evangelical	Total
Yes (users)	70%	30%	100%
No (non–users)	70%	30%	100%
Total	70%	30%	100%

Panel C: Does the percent using modern contraception differ by religious group?

(Answer: no)

Uses Modern Contraception:	Catholic	Evangelical	Total
Yes	40%	40%	40%
No	60%	60%	60%

Indicator

CONTINUATION RATES

Definition

The cumulative probability that acceptors of a contraceptive method will still be using any contraceptive method offered by the program after a specified period of time (e.g., one year). This is also known as the "all–method" continuation rate.

When using cross–sectional population data, the continuation rate for each interval of use (e.g., first, second, third, etc., month of use) is calculated as the complement of the ratio of acceptors who discontinue use of a program method of contraception at that duration to the number of women still using at the beginning of the month, i.e., 1 – discontinuation rate. These continuation rates are cumulated to obtain the probability that acceptors of a contraceptive method will still be using any program method after the specified period of time.

The indicator (CR_x) is calculated as:

$$CR_x = \prod_{x=1}^{x} (1 - q_x)$$

Where:

q_x = T_x / N_x = conditional probability of discontinuing use during a given interval (e.g., one month, one quarter, etc.);

T_x = the number of women discontinuing use during the interval;

N_x = number of women using at the beginning of the interval.

Note: Π signifies that $(1–q_x)$ is multiplied over all intervals from 1 to x.

Illustrative Computation

Continuation rates for modern methods for durations from 1–12 months, Indonesia, 1986–91.

x	T_x	N_x	q_x	CR_x
1	258.5	11839.7	.0218	97.8
2	137.5	11345.2	.0121	96.6
3	311.3	11047.0	.0282	93.9
4	106.3	10448.2	.0102	93.0
5	105.7	10304.8	.0103	92.0
6	264.4	10097.0	.0262	89.6
7	90.1	9586.7	.0094	88.7
8	71.2	9453.1	.0075	88.1
9	185.8	9263.9	.0201	86.3
10	99.4	8883.6	.0112	85.3
11	72.7	8707.7	.0084	84.6
12	280.1	8632.0	.0325	81.9

Source of data: Indonesia Demographic and Health Survey, 1991.

Note: numbers of discontinuations are not integers due to the application of sampling weights.

Data Requirements

Information on contraceptive initiation, use (including method switching), and discontinuation during a given reference period (e.g., the 3–5 years prior to a survey). Based on this information, one can calculate the percentage who have continuously used for a specific duration (e.g., 12 months, 18 months, etc.), as well as the median duration of use.

Data Source(s)

(Program–based): client records accompanied by a follow–up study of program drop–outs (see Chapter VII).

(Population–based): surveys with retrospective contraceptive use histories or calendars.

Purpose and Issues

As noted in Chapter VII, contraceptive continuation rates provide a useful summary measure of

the overall effectiveness of program services in enabling clients to sustain contraceptive use.

Although continuation rates can be calculated from either facility–based or population–based data, there are a number of limitations with facility–based data; thus, researchers are tending to look more to large scale surveys to provide more valid measurements of continuation among the target population.

The inclusion of contraceptive calendars in cross–sectional surveys such as the DHS minimizes the coverage problems discussed in Chapter VII. However, such surveys have limitations of their own. They (a) depend upon the accuracy of respondent recall, (b) do not allow linking of respondents to specific SDPs, and (c) may not capture the full contraceptive history (e.g., when five–year calendar is used). Such surveys do, however, allow the calculation of continuation rates for each specific method.

The two alternative continuation rates defined in Chapter VII (first–segment continuation rate and first–method continuation rate) apply to either program–based or population–based data.

It is important to note the distinction between discontinuation and failure of a contraceptive method. Discontinuation of contraception may occur because the individual chooses to stop using a selected method or because accidental pregnancy intervenes. As such, method failure is a subset of discontinuation. Method failure necessarily results in discontinuation. However, not all discontinuation is attributable to method failure (see the next indicator for further discussion of this point).

Indicator

USE FAILURE RATES

Definition

The cumulative probability that acceptors of a contraceptive method will experience an unintended pregnancy during use of a given contraceptive method within a specified period of time (e.g., one year).

When using cross–sectional population–based data on contraceptive use and failure, the failure rate for each interval of use (e.g., first, second, third, etc., month of use) is calculated as the ratio of unintended pregnancies occurring at that duration to the total number of users at the beginning of the month. The failure rates are cumulated to obtain the probability that acceptors of a contraceptive method will experience an unintended pregnancy within a specified period of time, such as one year.

"Use failure" includes conceptions due to method failure and user failure. Method failure refers to the theoretical probability of accidental pregnancy under the ideal condition of proper use by all women using a particular method. User failure refers to accidental pregnancies resulting from improper use.

The indicator (FR_x) is calculated as:

$$FR_x = 1 - [(1 - FR_{x-1}) - (CR_{x-1} * q_x)]$$

Where:

q_x = F_x / N_x = conditional probability of contraceptive failure during a given interval;

F_x = the number of women discontinuing use because of contraceptive failure during the interval;

N_x = number of women using at the beginning of the interval;

T_x = the number of women discontinuing use for all reasons, including contraceptive failure, during the interval;

$$CR_x = \prod_{x=1}^{x} [1 - (T_x / N_x)] =$$

the probability of acceptors continuing use after a specified period of time.

Illustrative Computation

Net use–failure rates, all methods, for durations from 1–12 months, Indonesia, 1986–91.

x	F_x	T_x	N_x	CR_x	q_x	FR_x
0	—	—	—	1.000	—	.0000
1	30.2	465.1	12656.7	.963	.0024	.0024
2	26.8	237.4	12100.6	.944	.0022	.0045
3	41.5	520.5	11766.3	.903	.0035	.0078
4	32.2	178.9	11135.0	.888	.0029	.0104
5	17.2	153.0	10965.5	.876	.0016	.0118
6	27.1	422.3	10746.0	.841	.0025	.0140
7	36.9	128.6	10211.6	.831	.0036	.0170
8	19.8	131.5	10043.0	.820	.0020	.0187
9	28.9	362.2	9833.6	.790	.0029	.0211
10	26.5	172.7	9433.1	.775	.0028	.0233
11	10.1	116.1	9230.5	.765	.0011	.0241
12	21.5	455.0	9136.5	.727	.0023	.0259

Source of data: Indonesia Demographic and Health Survey, 1991. Note: numbers of failures are not integers due to the application of sampling weights. Cumulative failure rates computed using increment–decrement life tables that take into account loss of exposure time resulting from discontinuation for reasons other than failure and censoring (see Namboodiri and Suchindram, 1987, for a discussion of increment–decrement life tables).

Data Requirements

The number of women of reproductive age who have experienced an unintended pregnancy while using a contraceptive method during a specified period of time by duration of use, the number of women discontinuing for other reasons by duration of use, and the number of women using contraception by duration of use. The specific information required depends upon the method used to calculate the measure (see below for further details).

Data Source(s)

Program–based: theoretically possible based upon follow–up surveys, but seldom calculated.

Population–based: surveys collecting retrospective information on contraceptive use and contraceptive–use status at the time of pregnancy for a specified reference period (for example, the 3–5 years prior to the survey).

Purpose and Issues

Use failure rates provide a measure of the relative frequency of occurrence of unintended pregnancies while a contraceptive method is being used. The indicator is viewed as a summary measure of the quality of contraceptive protection being provided by family planning programs, and in particular of program effectiveness in informing clients as to the proper use of methods. As noted above, the measure includes both method failures and failures resulting from improper use. As such, it is not possible to directly derive from the measure the relative importance of each cause of failure in a particular setting. However, since the theoretical failure rates for the major contraceptive methods are known from clinical trials, an unusually high use failure rate in a given setting can normally be attributed to high levels of improper method use.

A failure rate can be calculated either as the probability of a failure in the absence of all other reasons for discontinuation (gross rates) or as the probability of failure given that women may discontinue the method for some other reason (e.g., side effects, to become pregnant, etc.) and thus not be at risk of a failure (net rates). Net failure rates will always be lower than gross failure rates, since the higher the probability of discontinuing use of a method (for reasons other than failure), the lower the probability of discontinuing the method due to method failure. In general, net rates are more realistic because they capture the actual risk of women experiencing a contraceptive failure.

Use failure rates may be calculated from either program–based or population–based sources of data. In view of the limitations of program–based data discussed in Chapter VII and recent methodological advances, however, population–based sources are generally preferred (Jejeebhoy, 1989; Bongaarts and Rodriguez, 1991). The major limitations of estimates of use failure rates based upon program data are the same as those noted in the discussion of continuation rates.

Two types of failure rates have been used in prior studies: (a) the Pearl Pregnancy Rate (Pearl, 1932) and (b) life tables rates. Because of potential methodological problems with the Pearl Pregnancy Rate, most recent attempts to measure use failure have used life table methods.

Recent advances in demographic survey methodology and estimation techniques have made the calculation of use failure rates more feasible. One advance involves the collection and analysis of retrospective history data, especially the use of contraceptive–use calendars (Laing, 1985; Gaslonde and Carrasco, 1982). The illustrative computation shown above is based on contraceptive calendar data from the Indonesia DHS and life–table methods.

A second important advance has been the development of a “current–status” method to measure use failure (Bongaarts, 1984; Bongaarts and Rodriguez, 1991). This approach has the advantage over retrospective history or calendar methods of requiring considerably less detailed data. Under this method, only current contraceptive use status, the number of births during the two years prior to a survey, and the contraceptive use status at the time of pregnancy are required. The reader is referred to the sources cited above for further details on this method.

Chapter IX

Fertility Impact

A. Fertility Level

- Crude birth rate (CBR)
- Age–specific fertility rate (ASFR)
- Total fertility rate (TFR)

B. Births Averted

- Births averted (by the program)

C. Other Indicators

- Parity–specific birth rate
- Proportion of births above (or below) a specified parity
- Proportion of births by women above or below a specified age
- Median length of birth intervals
- Proportion of open or closed birth intervals that are of a specified length or longer
- Unwanted total fertility rate (UTFR)

Chapter IX

INDICATORS TO MEASURE FERTILITY IMPACT

This chapter presents a series of indicators for use in measuring family planning program impact on fertility. The selection of indicators presented was guided by two primary considerations. First, since a key concern in attempts to measure program fertility impact is the question of attribution of observed changes in fertility to the family planning program (as opposed to non–program factors), potential indicators were first screened in terms of their applicability in one or more accepted methods for addressing the issue of attribution (see below for further discussion). Secondly, as noted in the introduction to this Handbook, the focus of the initial version of the Handbook is on program impact on fertility. This scope will be expanded in the update of the Handbook to include in a more systematic fashion other areas of potential program impact (for example, on maternal and child health).

Considerable effort has gone into the development of methods and indicators for measuring the fertility or demographic impact of family planning programs. Reflecting this level of effort, at least eight distinct methodological approaches to quantifying the magnitude of program fertility impact may be found in the research literature (Sherris et al., 1985; United Nations, 1979, 1982, 1985, 1986; Chandrasekaran and Hermalin, 1975).

A review of the primary methods for measuring family planning program fertility impact reveals that two types of indicators are used in these methods: period fertility rates and/or births averted.

Period fertility rates (e.g., crude birth rate, total fertility rate, etc.) are used in two primary ways. A number of methods (e.g., the standardization/decomposition, trend analysis, experimental design, and areal regression methods) use period fertility rates measured at two or more points in time in order to estimate gross fertility change, and then attempt in varying ways to determine the portion of observed fertility change that is attributable to program contraception.

Other methods (e.g., multi–level methods, earlier applications of the areal regression method, and "after–only" experimental designs) use period fertility rates measured at one point in time and attempt to measure the extent to which observed cross–sectional differences in period fertility may be attributed to differences in levels of program exposure (e.g., program contraception) as opposed to other non–program determinants of fertility (e.g., socioeconomic differences and/or proximate determinants of fertility other than contraception).

Methods based upon births averted compare observed period fertility rates with the hypothetical rates that would have been expected in the absence of program contraception (referred to as "potential fertility"). The difference between observed and potential fertility is used to derive an estimate of the number of births that did not occur during a specified reference period due to program contraception. Included among these methods are the standard couple–years of protection, reproductive process analysis, component projection and prevalence methods. The reader is referred to the references cited above for computational details of these methods.

In the presentation of indicators, we begin by describing a series of conventional indicators that have been used in connection with one or more of the standard methods of fertility impact evaluation. The first section presents alternative indicators of fertility level, while the second section describes and illustrates the use of the births averted measure.

In the third section of the chapter, a series of additional indicators are presented under the heading of other indicators. In recent years, family planning program managers and researchers have

become increasingly interested in assessing the magnitude of program impact on aspects of fertility behavior that are not well measured by conventional fertility measures. One such area of interest is in indicators of the tempo of fertility, or birth spacing. In a number of countries in sub–Saharan Africa, for example, the demand for family planning for the purpose of spacing births is much stronger than that for limiting, and thus indicators relating to the spacing or timing of births are relevant short–run impact indicators for programs in such countries.

Another area of interest is indicators for use in programs with explicit parity and maternal–age related targets; for example, programs with objectives to reduce the number or rate of births to women above a certain parity (e.g., parity 5) or to women below/above certain ages (e.g., below age 20 or above age 35). Conventional fertility measures may not be sensitive indicators of impact in the context of such programs.

Other relevant areas include program impact on reducing the extent of unwanted fertility and the incidence of abortion.

While the types of indicators described in the preceding paragraphs have not heretofore been widely used in evaluating impact, the fact that fertility impact targets for a sizeable number of programs in contemporary developing country settings are stated in parity– or maternal age–specific terms requires that some consideration be given to appropriate impact indicators for such programs in this Handbook. It should be noted, however, that we have limited the indicators considered to those that are felt to be: (1) sufficiently sensitive and specific, and (2) based upon reasonably accurate data in a large number of developing countries.

In some cases, we have not proposed an indicator for dimensions of fertility outcomes that are viewed as important/relevant due to concerns over data availability and quality. For example, while abortion rates are viewed as a valid indicator of impact for some family planning programs, the lack of accurate data on abortions in most countries and the unproven validity of alternative data collection and (indirect) estimation approaches have led us not to include an indicator on the topic of abortion in the current edition of the Handbook.

Finally, it should be noted that relatively little work has been done on the development of impact evaluation methods that address the question of program attribution using indicators of the type described in the third section of this chapter. The EVALUATION Project hopes to explore the use of non–traditional fertility indicators for program evaluation purposes over the course of the project, and the update of this Handbook will reflect progress made in this area.

Indicator

CRUDE BIRTH RATE (CBR)

Definition

The number of births occurring in a given year or other specified reference period per 1,000 population.

The CBR is calculated as:

$$CBR = (B / P) * 1{,}000$$

where:

B = number of births in a given year or reference period, and

P = total mid–year or mid–period population.

Illustrative Computation

Estimate of the CBR in Ghana, 1985–1988.

$$CBR = (B / P) * 1{,}000$$

$$= (575{,}381 / 13{,}667{,}229) * 1{,}000$$

$$= 42.1$$

Source of data:

Births: Ghana Demographic and Health Survey, 1988. (Note – the figure shown was derived by applying survey sampling weights to the reported numbers of births during the three–year reference period, adjusting for household– and individual–level survey non–response, and dividing by the number of years in the reference period. See Appendix B of the Ghana DHS Report for details on sampling weights and survey non–response.)

Population: estimated total population of Ghana for January 1, 1987 (the mid–point of the reference period for the indicator), U.S. Bureau of the Census, Center for International Research.

Data Requirements

The total number of births occurring in a given year or reference period.

The enumerated or estimated total mid–year or mid–period population for the same year/period.

Data Source(s)

Births: vital statistics, population censuses or population–based surveys.

Population: normally obtained population census data for census years and from population estimates for off–census years, but can also be estimated from population–based surveys under certain conditions (see below).

Purposes and Issues

Sometimes referred to simply as "the birth rate," the CBR is among the least demanding measures of period or current fertility in terms of data requirements. Annual counts of births are available from vital statistics in most countries. In many countries, however, vital registration systems are plagued by serious levels of under–registration. Alternatively, estimates of the number of births occurring during specified periods can be derived from answers to census or survey questions, or from partial or complete birth histories in sample surveys.

While the CBR is normally expressed as an annual rate, longer reference periods are often used for computational purposes when calculated from survey data; for example, the 3–5 years prior to the survey. The reasons for this are twofold: (1) to increase the numbers of events reported as a means of dampening sampling variability in the estimates of numbers of births from year to year, and (b) to dampen the effects of reference period errors in the reporting of ages/dates of birth of children associated with the preference for digits that are multiples of 12. In such instances, the average annual CBR for the period covered is normally calculated (as in the illustrative computation shown above).

In addition to official population estimates or projections, an estimate of the total population for the denominator of the CBR may also be obtained by applying survey sampling weights to the counts of household members from the household roster portion of surveys with unrestricted population universes; that is, from surveys covering all households as opposed to only households containing at least one female of reproductive age, a universe

restriction imposed in some demographic surveys. Unless the proportion of the population that is excluded from surveys with restricted universe definitions is known or can be estimated, such surveys will under–estimate the total population and thus upwardly bias estimates of the CBR.

It should be noted that since the CBR does not take into account the age–sex composition of the population, comparisons between populations and assessment of trends over time using the CBR may be distorted by differences or changes in population composition. For this reason, inferences regarding program fertility impact based upon the CBR should be made with caution. At a minimum, age–standardization should be employed when making comparisons or assessing trends based upon the CBR. More powerful methods for dealing with the influences of compositional differences and other potentially confounding factors are described in the references cited in the introduction to this chapter.

Indicator

AGE–SPECIFIC FERTILITY RATE (ASFR)

Definition

The number of births occurring during a given year or reference period per 1,000 women of reproductive age classified in single–or five–year age groups.

The ASFR is calculated as:

$$ASFR_a = (B_a / E_a) * 1{,}000$$

where:

B_a = number of births to women in age group *a* in a given year or reference period, and

E_a = number of person–years of exposure in age group *a* during the specified reference period.

Illustrative Computation

Estimates of average annual ASFRs for all women of reproductive age, Ghana, 1985–88 period.

Age Group	Births B_a	Person-Years of Exposure E_a	Rate/ Woman	Rate/1,000 Person-Years
15-19	351	2810.8	.125	125
20-24	747	2864.7	.261	261
25-29	743	2658.7	.279	279
30-34	473	1912.9	.247	247
35-39	287	1490.2	.193	193
40-44	124	1056.2	.117	117
45-49	44	730.1	.060	060

Source of data: Ghana Demographic and Health Survey, 1988.

Data Requirements

The number of births in a given year or reference period classified by age of mother.

The number of women of reproductive age (i.e., 15–44 or 15–49 years), in 1– or 5–year age groups.

Data Source(s)

Vital statistics (numerator only), population censuses or population–based surveys.

Purposes and Issues

The ASFR has two primary uses: (1) as a measure of the age pattern of fertility, that is of the relative frequency of childbearing among women of different ages within the reproductive years, and (2) as an intermediate computation in the derivation of the Total Fertility Rate (TFR), discussed next in this chapter.

As indicated above, ASFRs may be derived from several sources. When counts of births are derived from vital registration, population projections or estimates of the number of women in each age group within the reproductive years are required for the denominator of the rate. When derived from population censuses or surveys, both the numerator and denominator of the rate may be derived from the census or survey. Estimates from censuses are derived from questions on births during a specified period preceding the census (usually 12 months), while survey estimates may be derived either from questions on births within a specified prior period or from partial or complete birth histories.

A simpler, although less precise, procedure for computing the denominator of the rate is to take the average of the number of women in each age group during the reference period covered by the measure (i.e., the average of the numbers of women in each age group at the beginning and end of the reference period).

Reference periods of more than one year are frequently used in the computation of ASFRs from survey data, the rationale being the same as described for the CBR: to dampen sampling variability associated with relatively small numbers of annual births occurring to women in single or five year age groups and the distorting effects of reference period reporting errors. Various analyses of DHS fertility data, for example, alternately use the 3 or 5 years period prior to the survey in calculating ASFRs (Arnold and Blanc, 1989; Lutz, 1990). When multiple years are used for computational purposes, average annual rates are normally presented.

Unlike the CBR, the ASFR is unaffected by differences or changes in population age composition, and thus is more useful in comparing different populations or sub–groups and in measuring changes over time. The ASFR is, however, affected by differences or changes in the number or proportion of women exposed to the risk of pregnancy. Thus, changes in ASFRs may provide misleading information regarding the impact of family planning programs on fertility when other factors affecting risk of pregnancy are changing; for example, for the 15–19 and 20–24 age groups when age at marriage is rising quickly.

To address this problem, ASFRs may be calculated only for women who were continuously married or in union during the reference period of the measure. The resulting measure is known as the Marital Age–Specific Fertility Rate (MASFR). However, in order to calculate this measure, data on duration of marriage or marriage histories are required. In actual practice, MASFRs are more often approximated by calculating ASFRs for women married or in union at the time of a survey, although it should be recognized that this is only an approximation of the MASFR since women who are married or in union at the time of a given survey may not have been continuously married or in union over the entire reference period of the measure (e.g., for the 3–5 years prior to the survey).

ASFRs are sometimes presented for different groups of women; for example, ASFRs are for women currently married or in union and for all women of reproductive age in DHS country reports. In societies where fertility is largely confined to marriage, ASFRs for women currently married or in union will provide more or less complete coverage of recent fertility. Where a large share of fertility occurs outside of recognized unions, however, the restriction of the ASFR to currently married women will result in an under–estimate of the level of current fertility.

Indicator

TOTAL FERTILITY RATE (TFR)

Definition

The number of children that would be born per woman (or per 1,000 women) if she/they were to pass through the childbearing years bearing children according to a current schedule of age–specific fertility rates.

The TFR is calculated as:

TFR = Σ $ASFR_a$ (for single–year age groups), or

TFR = 5 Σ $ASFR_a$ (for 5–year age groups),

where:

$ASFR_a$= age–specific fertility rate for women in age group *a* (expressed as a rate per woman).

Illustrative Computation

Estimate of the average annual TFR for all women of reproductive age in Ghana, 1985–88 period.

TFR = 5 (.125 + .261 + .279 + .247 + .193 + .117 + .060)

= 6.11

Where: the figures in parentheses are age–specific rates for the 15–19, 20–24, ... , 45–49 age categories, respectively.

Source of data: Ghana Demographic and Health Survey, 1988.

Data Requirements

A current schedule of age–specific fertility rates (ASFRs), for 1– or 5–year age groups.

Data Source(s)

Vital statistics (numerator only), population censuses or population–based surveys.

Purposes and Issues

The TFR is the most widely–used fertility measure in program impact evaluations for two main reasons: (1) it is unaffected by differences or changes in age–sex composition, and (2) it provides an easily understandable measure of hypothetical completed fertility.

Although derived from the ASFR, a period fertility rate, the TFR is a measure of the anticipated level of completed fertility per woman (or per 1,000 women) if she/they were to pass through the reproductive years bearing children according to the current schedule of ASFRs. It should be emphasized that the TFR is only a hypothetical measure of completed fertility, and it is thus possible that women of reproductive age at any given point in time may have completed family sizes that are considerably different from that implied by a current TFR, should age–specific fertility rates rise or fall in the future.

As the TFR is derived from a schedule of ASFRs, the comments and caveats regarding the ASFR also apply to the TFR; i.e., method of computation from different sources of data, effects of changing exposure to pregnancy, and implications of computation for currently married versus all women of reproductive age. As was also the case for the ASFR, the TFR may be computed for women who were continuously married or in union during the reference period of the measure in order to dampen the potentially confounding effects of differences in exposure to the risk of pregnancy (to the extent that these are associated with marital status). This measure is known as the Total Marital Fertility Rate (TMFR). See the discussion of the Marital Age–Specific Fertility Rate (MASFR) for details on data requirements for the computation of this measure.

Note also that while the standard age range for the TFR is ages 15–49, TFRs for other age ranges (e.g., 15–34) are sometimes used for analytic purposes; for example, in order to dampen the influences of truncation when examining cohort trends from birth history data (the reader is referred to any DHS country report for an illustration of this).

Indicator

BIRTHS AVERTED (BY THE PROGRAM)

Definition

The number of births that did not occur during a specified reference period due to program effects (services and attitudinal changes).

The number of births averted during a specified period is equal to the difference between potential and observed fertility during the period. Potential fertility is defined as the fertility level that would have prevailed during the reference period in the absence of the family planning program.

Data Requirements

Note: data requirements and sources of data are indicated below for the method most often used to derive estimates of births averted: the prevalence methods. The reader is referred to the series of United Nations manuals on methods for measuring the fertility impact of family planning programs for computational details on this and other methods that may be used to calculate births averted (United Nations, 1979, 1982, 1985, 1986).

Prevalence Method:

- estimates of the contraceptive prevalence rate (CPR) among currently married women (see Chapter VIII for computational details) and ASFRs for a common reference period;
- the number of women of reproductive age in five–year age groups;
- the enumerated or estimated total population size; and
- data on continuation rates and use–effectiveness (optional: if country–specific data are not available, standard values may be used). See the references cited above for further details.

Data Source(s)

Population–based surveys.

Purposes and Issues

The births averted indicator provides a conceptually appealing measure of the fertility impact of family planning programs – the number of births that did not occur during a given reference period as the result of family planning program efforts. The measure is based upon a comparison between observed fertility during a given reference period and "potential fertility," the level of fertility that would be expected in the absence of the family planning program.

Each of the principal methods for measuring program fertility impact enumerated in the introduction to this chapter may be used to derive estimates of births averted. The primary differences among the various methods lie in the underlying assumptions and procedures used in calculating potential fertility. Potential fertility may be estimated from: (1) previous fertility levels of program clients, (2) fertility levels in the general population under study, or (3) estimates of "natural fertility," the fertility level that would prevail in the population under study in the absence of any contraception (program or non–program).

A number of complications arise in the calculation of potential fertility. One issue concerns the treatment of substitution of non–program for program sources of supply. That is, it is necessary to account for the fact that some contraceptive users relying on program sources would have used or would shift to non–program sources of supply in the absence of the program, thus resulting in lower fertility than would prevail if there were no alternative sources of supply available. Thus, a distinction needs to be made between gross potential fertility, which is defined as the level of fertility that would prevail if all use of program contraception were eliminated and there were no switching to non–program sources (or no alternative sources of supply available), and net potential fertility, which is defined as the level of fertility that would prevail if substitution of non–program for program sources is taken into account. In most instances, net potential fertility will fall somewhere between observed fertility levels and gross potential fertility, as there are normally alternative sources of supply available in most settings (Bongaarts, 1985). The births averted by a family planning program measure thus must take into account the fact that all births averted due to contraception during a particular period may not be attributable to the program.

Another issue concerns the possible "catalytic" effect of programs; that is, the extent to which program I–E–C and related efforts affect the fertility of women and couples who are not program clients; for example, through attitudinal changes and/or use of non–program contraception.

Each of the approaches to calculating potential fertility is problematic in certain respects, and thus estimates of births averted should be interpreted in light of the assumptions underlying the method used to derive estimates of potential fertility. The series of United Nations manuals referenced above should be consulted for details on the assumptions and limitations involved in each method for calculating births averted.

Among the alternative methods for computing births averted, the Prevalence Method would appear to have an advantage in that it requires only estimates of contraceptive prevalence (preferably by program/non–program source of supply) and marital age–specific fertility rates in settings where the model parameters of the method are reasonable (which would appear to cover most practical situations). It should be noted, however, that it is not always easy to separate program from non–program contraception in survey data, since some users may not know if their source of supply was program or non–program.

An illustrative computation of births averted using the Prevalence Method is provided on the right.

Illustrative Computation

Age–specific estimates of births averted by the Tunisian family planning program, 1978–79, using the Prevalence Method.

$BA_a = (PAF_a - AF_a)\ POP_a$

$BAN_a = (NAF_a - PAF_a)\ POP_a$

Where:

BA_a = births averted among women currently married or in union in age group *a* due to program contraception,

BAN_a = births averted among women currently married or in union in age group *a* due to non–program contraception,

NAF_a = the estimated level of natural fertility; $NAF_a = AF_a/[1-C_a(u'_a+u''_a)]$,

PAF_a = the estimated level of potential fertility $PAF_a = AF_a(1-C_a \times u''_a)/[1-C_a(u'_a+u''_a)]$,

u'_a = prevalence of program contraception among women currently married or in union in age group *a*,

u''_a = prevalence of non–program contraception among women currently married or in union in age group *a*,

AF_a = marital age–specific fertility rate for age group *a*,

POP_a = the number of women in age group *a*, and

C_a = elasticity coefficient for age group *a* (see standard schedule below).

Age Group	Coefficient
15–19	0.620
20–24	0.620
25–29	0.823
30–34	0.940
35–39	1.022
40–44	1.309
45–49	1.89 8

Source of data: Bongaarts (185:101)

Prevalence of Program and Non–program Contraception, Observed Fertility Rates and Estimated Natural and Gross Potential Fertility Rates, by Age Group, Tunisia

(Prevalence in percentage; fertility rates per 1,000)

Age Group a	Prevalence of: Program Contraception 1978 u'_a	Prevalence of: Non–program Contraception 1978 u''_a	Observed Fertility Rate 1979 AF_a	Natural Fertility Rate 1979 NAF_a	Gross Potential Fertility Rate 1979 PAF_a
15–19	4	2	35.5	36.9	36.4
20–24	8	4	216.2	233.6	227.8
25–29	15	7	273.8	334.3	315.1
30–34	24	7	235.4	332.2	310.4
35–39	28	5	190.3	287.1	272.5
40–44	27	4	85.0	143.1	135.6
45–49	20	1	23.4	38.9	38.2

Gross Fertility Effects and Gross Births Averted by Program and Non–program Contraception, by Age Group, Tunisia, 1979

Age Group a	Female Population,1979 POP_a (thousands)	Gross Fertility Effect of: Program Contraception $PAF_a–AF_a$ (per 1,000)	Gross Fertility Effect of: Non–program Contraception $NAF_a–PAF_a$ (per 1,000)	Gross Births Averted by: Program Contraception BA_a	Gross Births Averted by: Non–program Contraception BAN_a
15–19	347.6	10	5	313	171
20–24	292.2	23	11	3,390	1,695
25–29	218.5	53	25	9,024	4,195
30–34	145.1	95	28	10,883	3,163
35–39	154.7	89	16	12,716	2,259
40–44	149.8	65	10	7,580	1,124
45–49	130.3	17	1	1,929	91
TOTAL	–	–	–	45,835	12,701

Source of data: adapted from Jemai (1985)

Indicator

PARITY–SPECIFIC BIRTH RATE

Definition

The number of births of parity "i" in a given year or reference period per 1,000 women of reproductive age.

The parity–specific birth is calculated as follows:

$$PSBRi = (B_i / E_{f,15–49}) * 1,000$$

where:

$PSBR_i$ = the rate per 1,000 women of reproductive age of births of parity *i*,

B_i = the number of births of parity *i* in a given year or reference period, and

$E_{f,15–49}$ = the number of person–years lived during the reference period by women 15–49 years of age.

Illustrative Computation

Parity–specific birth rates, Ghana, 1985–88 period.

Births Parity	B_i	Person–Years of Exposure E_a	Rate/1000 Person–Years
1	524	13523.6	38.9
2	485	13523.6	35.9
3	437	13523.6	32.3
4	334	13523.6	24.7
5	319	13523.6	23.6
6	238	13523.6	17.6
7	160	13523.6	11.8
8+	272	13523.6	20.1
Total	2,769		

Source of data: Ghana Demographic and Health Survey, 1988.

Data Requirements

The total number of births during a given reference period classified by parity.

The average number of women of reproductive age during the reference period.

Data Source(s)

Vital statistics (if information on parity is recorded), population censuses, or population–based surveys.

Purposes and Issues

The parity–specific birth rate provides a measure of differential fertility by parity. Its principal use is in examining the relative frequency of births of different parities and changes in frequencies over time. The distribution of births by parity is of interest for program monitoring and evaluation purposes for several reasons. First, since decisions to continue childbearing or to regulate fertility are very often parity–specific, it is likely that "unwanted" births will tend to be concentrated among higher parity births (see Chapter V for definitions of "wanted" and "unwanted" births).

Secondly, higher parity (along with primiparity) births are subject to higher maternal–child health risks than births of lower orders (National Research Council, 1989). Thus, declines in birth rates for higher parities provide an indication of family planning program success in enabling women or couples to better control unwanted fertility and/or in minimizing the incidence of "high risk" pregnancies (on grounds of high parity).

Thirdly, some researchers have suggested that declining fertility brought about by increased use of contraception might be more evident from parity–specific fertility measures than from measures that include all women of reproductive age (as, for example, in the TFR) (Ryder, 1982; Srinivasan and Freymann, 1989). Thus, it may be the case that parity–specific measures are more responsive or sensitive to changing fertility than conventional measures of fertility level, although further research is needed on this topic.

Finally, the indicator is also useful for monitoring and evaluation purposes in programs that

have explicit parity–based objectives; for example, in programs with an explicit objective of reducing the incidence of pregnancies among high–parity women (e.g., parity 5 and above) or with explicit "ideal family size" targets (e.g., 2–child families).

Sources of data and modes of questioning in censuses and surveys for this indicator are essentially the same as for the age–specific fertility rate, with the added requirement that the parity of births occurring during the specified reference period be either recorded in vital registration and facility–based record systems or reported on censuses and surveys. When based upon registration or facility–based data, it should be recognized that since not all births are recorded in such systems, measures based upon these sources of data may not provide an accurate picture of trends in the larger population.

More sophisticated versions of parity–based measures have as their denominator either the number of women at parity (i–1) at the beginning of the specified reference period or the number of person–years lived at parity (i–1) during the interval, and thus provide a measure of the probability of making the transition from one parity to the next (i.e., from parity i–1 to parity i) during the specified period; for example, the parity progression ratio (Lutz, 1990; Ryder, 1982). These measures are not presented here because they (1) require quite large samples in order to derive reliable estimates and (2) require the application of analytic methods (e.g., life table methods) that are beyond the scope of this Handbook.

Indicator

Proportion of Births Above (or Below) a Specified Parity

Definition

The proportion of births occurring during a given reference period that are above (or below) a specified parity (e.g., the proportion of births that are parity 5 or higher).

The indicator is calculated as:

$$ORDGE_{(i)} = B_{GE(i)} / B$$

Where:

$ORDGE_{(i)}$ = the proportion of births during a given reference period of parity *i* or above,

$B_{GE(i)}$ = the number of births of parity *i* or above, and

B = the total number of births during the reference period (all parities).

Note: the proportion of births below a specified parity is calculated in a similar fashion, substituting the number of births below the specified parity, $BLE_{(i)}$, for $BGE_{(i)}$.

Illustrative Computation

Proportion of births during the 1985–88 period in Ghana that are parity 5 and above (data derived from the distribution of births by parity displayed in the illustrative computation for the preceding indicator).

$$ORDGE_{(i)} = B_{GE(i)} / B$$
$$= 989 / 2{,}769 = 35.7$$

Source of data: Ghana Demographic and Health Survey, 1988.

Data Requirements

The total number of births during a given year or reference period classified by parity.

Data Source(s)

Vital statistics (if information on parity is recorded), facility–based data (e.g., data on the parities of attended births), or population–based surveys.

Purposes and Issues

This indicator has been proposed as an alternative parity–based fertility measure. An advantage of this specification is its simpler interpretation; that is, as a single number instead of a series of parity–specific rates.

The indicator has been used in several different ways for program monitoring and evaluation purposes. A primary use is as an indicator of the program success in reducing the relative frequency of births that are above a program target parity. The primary rationale for the indicator's use in this fashion is that pregnancies above parity five have been shown to be associated with increased maternal–child health risks (National Research Council, 1989). Thus, the indicator provides a means of monitoring program success in enabling women and couples to avoid pregnancies that threaten the health of mothers and children.

Other researchers have used the indicator in a more "demographic" manner; that is, as a proxy measure for conventional population–based measures of fertility level (e.g., the TFR). The rationale for the indicator's use in this manner is that declining fertility should be reflected by declining proportions of higher–order births or, conversely, by increasing proportions of low order births. These expectations are borne out in two recent studies. A study in Thailand, for example, demonstrates strong negative province–level correlations between the proportion of registered births that are of orders 1 or 2 and the CBR, TFR and CPR (Prasartkul et al., 1987). Similarly, a recent three–country study based on survey data demonstrates a high correlation between the proportion of births of order 5 or higher and the TFR (Srinivasan et al., 1992). It should be recognized, however, that it is possible to observe a decline in the proportion of higher–order births simply because of a temporary rise in the number of lower–order births associated with changes in age distribution or marital patterns. Accordingly, the indicator should be interpreted with these caveats in mind.

As noted earlier, some researchers have argued that parity–based indicators may be more sensitive to short–term changes in fertility behavior

than conventional fertility measures, although further research is needed on this issue (Ryder, 1982; Srinivasan and Freymann, 1989).

The indicator may be derived from the same sources of data as conventional period fertility measures. Computation from vital registration data requires, however, that the birth order of registered births be recorded on the birth records.

Since the indicator may be meaningfully computed from registration and facility–based data, it may be used for monitoring purposes in situations where reliable population counts or estimates for the denominators of conventional fertility measures are not available (i.e., in "numerator–only" analyses). This feature is especially useful in monitoring trends for small geographic areas. Caution should be exercised, however, in the interpretation of such data since they pertain only to women whose births are captured in facility–based data collection systems. To the extent that these women are selective of women giving birth in a particular year, conclusions regarding levels and trends in the indicator may not be valid for national populations.

Indicator

Proportion of Births by Women Above or Below a Specified Age

Definition

The proportion of births during a given year or other reference period that are to women above or below specified age (e.g., above age 35 or below age 20).

The indicator is calculated as:

$$PBLT20 = B_{LT20} / B$$

Where:

PBLT20 = the proportion of births during a given year or reference period to women less than 20 years of age,

B_{LT20} = the number of births during the specified reference period to women less than 20 years of age, and

B = the total number of births during the given reference period.

Note: the indicator for births above age 35 would be calculated in an analogous fashion, but substituting the number of births to women aged 35 years and above in the numerator of the measure.

Illustrative Computation

Proportion of births during the 1985–88 period to women less than 20 years of age and 35 or more years of age, Ghana.

$PBLT20 = B_{LT20} / B$	$PBGE35 = B_{GE35} / B$
= 351 / 2,769	= 455 / 2,679
= 12.7%	= 16.4%

Source of data: Ghana Demographic and Health Survey, 1988.

Data Requirements

The total number of births during a given year or reference period, classified by maternal age.

Data Source(s)

Vital statistics (where information on maternal age is recorded), facility–based data (e.g., delivery records), population censuses, or population–based surveys.

Purposes and Issues

This indicator provides a measure of the relative frequency of births during a given reference period to women who are viewed as "too young" or "too old" for childbearing. The concern over maternal age has several dimensions. One concern is that bearing children under age 20 (some would place the "cut–off" at age 18 or 35 and above) places women and infants at elevated health risks (National Research Council, 1989). For program monitoring and evaluation purposes, this indicator provides a measure of success in both motivating women to avoid high risk pregnancies on the grounds of maternal age and in providing the necessary contraceptive services to enable them to do so.

A second concern is more demographic in nature; there is considerable evidence that young mean age at first birth is strongly associated with higher fertility. Thus, decreasing proportions of births to young women implies rising ages at first birth.

In addition, a recent study demonstrates the utility of an alternative specification of the indicator, the proportion of births that are to women less than 35 years of age, as a simple measure of the extent of fertility control in national populations (Anderson and Silver, 1992).

Because the required information is available from client records in many maternal–child health programs (also from vital statistics, population censuses, and sample surveys) the indicator has seen increasing use in recent years. The indicator is especially relevant in family planning programs with explicit maternal–child health objectives. The indicator is sometimes used in constructing a composite indicator of high–risk births, taking into account maternal age, parity, and birth interval lengths.

Since the indicator may be computed from facility–based data, it may be used in "numerator–only" analyses, subject to the limitations of such indicators noted earlier. As was also the case with the previous indicator, it should be recognized that declines in the proportion of births among younger (or older) women might result from increases in the number of births among women in other age categories. This should be borne in mind when interpreting the indicator.

Indicator

MEDIAN LENGTH OF BIRTH INTERVALS

Definition

Open interval measure: the median number of months between a specified reference date (for example, the date of a survey) and the last birth among women with one or more births.

Closed interval measure: thºe median number of months separating successive births among women with two or more births.

Both indicators are calculated as:

MEDIAN = L + [(50 – cf / f) * i]

Where:

L = the true lower limit of the class interval in which the median is located,

50 = the 50 percentile observation,

cf = the cumulated frequency up to the median class interval,

f = the frequency within the median class interval, and

i = the class width.

Note: the above formula is for use with grouped data consisting of percentage frequencies in each class. For ungrouped data, the median is the value of the observation falling at exactly the 50th percentile of the distribution of observations.

Illustrative Computation

Estimate of median length of open birth intervals among women currently married or in union, Ghana, 1988.

Months Since Last Birth	f (percent)	cf (percent)
7–17	7.1	7.1
18–23	11.0	18.1
24–36	39.4	57.5
37–47	20.9	78.4
48+	21.6	100.0

MEDIAN = L + [(50 – cf / f) * i]

= 24 + [(50 – 18.1 / 39.4) * 12]

= 33.7

Source: Ghana Demographic and Health Survey, 1988.

Data Requirements

Open interval measure: the distribution of women currently married or in union with one or more births by number of months since their last birth.

Closed interval measure: the distribution of births occurring during a specified reference period (e.g., the 3–5 years prior to a survey) occurring to women with 2 or more births by number of months between successive births.

Data Source(s):

Open interval measure: population–based surveys.

Closed interval measure: vital statistics or facility–based data (where the date of the birth preceding the registered birth is recorded), or population–based surveys.

Purposes and Issues

Birth interval measures provide information on a dimension of fertility not addressed by measures of fertility level: the tempo or spacing of births. Measures of birth spacing have several important uses in program monitoring and evaluation efforts. First, proper spacing of births is an essential element of family planning programs with priority maternal–child health objectives, and thus indicators are required to monitor and evaluate program performance in achieving such objectives.

Secondly, in societies where a large share of demand for family planning is for spacing as opposed to limiting (for example, in a number of countries of sub–Saharan Africa), birth interval measures may provide an indication of short–run program success in enabling women and couples to implement their reproductive preferences (see Chapter VI for details on the measurement of demand for family planning).

Thirdly, some researchers have suggested that changes in the length of birth intervals may provide a more sensitive indicator of changing fertility behavior than conventional fertility measures such as the TFR (Ryder, 1982; Srinivasan and Freymann, 1989). Further research is needed, however, on this issue.

In addition to issues of spacing, it should be recognized that the pace of fertility, as reflected in the length of intervals separating births, influences the rate of population growth independently of changes in levels of completed fertility through its effect on the length of generations. However, in most instances longer birth intervals translate into lower completed fertility due to "lost" exposure time during the peak reproductive years.

Caution should be exercised in the interpretation of closed birth interval lengths over extended periods of time, since over the long run birth interval lengths may not be highly correlated with levels of completed fertility. In the more developed countries, for example, childbearing is often confined to a small portion of women's reproductive years, with median birth intervals being relatively short (2–3 years, for example). In these societies, women and couples tend to be highly successful in avoiding additional pregnancies once the desired number of children is reached, and thus low fertility levels are observed concurrently with relatively short birth intervals. Thus, understanding the dynamics and implications of changes in closed birth interval measures over time requires that information on trends in contraceptive use for limiting and spacing purposes be considered.

Open birth interval data should be limited, to the extent possible, to fecund women, since sizeable numbers of infecund women can distort estimates of median interval lengths through their disproportionate contribution of long intervals. In most instances, open birth interval data are limited to women currently married or in union in order to restrict the universe of the indicator to women presumably fecund and at risk of pregnancy. As noted earlier, however, such a restriction will overstate the number of women who are fecund, and it will understate the number of women actually at risk of pregnancy in populations where a sizeable share of childbearing occurs outside of marriage.

Indicator

Proportion of Open or Closed Birth Intervals That Are of a Specified Length or Longer

Definition

Open interval measure: the proportion of women currently married or in union whose open birth interval, that is, the interval between the date of the last birth and a reference date (e.g., the date of a survey), is of a specified length or longer; for example, 24 months or longer.

Closed interval measure: the proportion of closed birth intervals, that is, the length of the period separating successive births, that are of a specified length or longer; for example, 24 months or longer.

The open interval measure is calculated as:

$$OBIGE_{(i)} = P_{OBIGE(i)} / P$$

Where:

$OBIGE_{(i)}$ = the proportion of women with an open birth interval of length *i* months or longer,

$P_{OBIGE(i)}$ = the number of women with an open birth interval of length *i* months or longer, and

P = the total number of women.

Note: The closed interval measure is calculated in a similar fashion, substituting closed for open intervals.

Illustrative Computation

Proportion of open birth intervals of length 24 months or longer among women currently married or in union, Ghana, 1988.

$$OBIGE_{(24)} = P_{OBIGE(24)} / P$$
$$= 2{,}711 / 3{,}310 = 81.9\%$$

Source of data: Ghana Demographic and Health Survey, 1988 (see the tabulation of open birth interval lengths presented for the previous indicator).

Data Requirements

Open interval measure: the distribution of women currently married or in union with one or more births by number of months since the last birth.

Closed interval measure: the distribution of births occurring during a specified reference period (e.g., the 3–5 years prior to a survey) by the number of months separating the birth and the previous birth among women with two or more births.

Data Source(s)

Open interval measure: population–based surveys.

Closed interval measure: vital statistics and facility–based data (where the date of the birth preceding the registered birth is recorded), or population–based surveys.

Purposes and Issues

This indicator has been proposed as an alternative to median birth interval lengths as a measure of the tempo or spacing of births. Its principal advantages are simpler interpretation and its utility in numerator–only analyses.

A principal use of the present indicator is in family planning programs with explicit birth–spacing objectives; for example, programs with a primary objective of reducing the incidence of short–interval births (i.e., those less than 18 or 24 months). There is considerable empirical evidence linking short birth intervals to elevated risk for infant mortality (National Research Council, 1989). For program evaluation purposes, the indicator provides a measure of the extent of program success in minimizing the incidence of short–interval births through post–partum contraception.

In addition to this use, a recent cross–national study provides empirical evidence of an association between the present indicator and conventional period fertility measures (Srinivasan et al., 1992). In the study, a high inverse correlation (approximately .86) was observed between the proportion of open birth intervals of length 24 months or longer and the TFR. A somewhat lower, but still highly significant, correlation was found in the same study between the proportion of last closed birth intervals that were 24 months duration or greater and the TFR.

When birth history data from sample surveys are available, an alternative specification of the closed birth history measure would be as the proportion of all birth intervals closed during a specified reference period (e.g., the 3–5 years

Indicator

UNWANTED TOTAL FERTILITY RATE (UTFR)

Definition

The number of unwanted children that would be born per woman (or per 1,000 women) if she/they were to pass through the reproductive years bearing children according to current schedules of unwanted fertility.

For this indicator, "unwanted" births are defined as the complement of "wanted" births as defined in connection with the wanted total fertility rate (WTFR) (see Chapter V).

The indicator is calculated as:

$$UTFR = 5\,(B_{a,u} / E_a), \text{ or}$$

$$= TFR - WTFR$$

Where:

$B_{a,u}$ = the number of births to women in age group *a* during a given year or reference period that are unwanted,

E_a = the number of person–years lived by women in age group *a* during the reference period,

TFR = the total fertility rate for a given year or reference period, and

WTFR = the wanted total fertility rate (see Chapter V for computational details).

Illustrative Computation

Estimate of the UTFR for women aged 15–44 years in the 36 months prior to the survey, Northeast Brazil, 1991.

$$UTFR = TFR - WTFR$$

$$= 3.66 - 2.13$$

$$= 1.53$$

Source of data: Brazil Demographic and Health Survey, 1991. See Chapter V for computational details on the WTFR generally and the computations for Northeast Brazil specifically.

Data Requirements

Responses to survey questions, by age of woman, on:

- numbers and dates of births during a recent period (typically the 2–5 years prior to a survey);
- desired number of children or family size (see Chapter V);
- number of surviving children; and
- desire for additional children (see Chapter V).

Data Source(s)

Population–based surveys.

Comments and Caveats

The UTFR provides a hypothetical measure of the average number of "unwanted" births a woman or cohort of women would have during her/their reproductive career(s) if they were to follow current schedules of unwanted fertility. As the wanted status of births is based upon reproductive preferences or demand for children (see Chapter V), the indicator provides a conceptually direct measure of family planning program impact in enabling women and couples to achieve their reproductive goals (i.e., to avoid unwanted pregnancies).

In the illustrative computation for Brazil, for example, the estimate of 1.5 indicates that women in Northeast Brazil would have on average 1.5 unwanted births over the course of their reproductive years if current levels of age–specific fertility and demand for children were to prevail throughout the reproductive years of women of reproductive age at the time of the 1991 DHS, suggesting the existence of a significant level of failure to control fertility so as to satisfy fertility preferences.

Unwanted births in the present indicator are the arithmetic complement of "wanted" births as defined in Chapter V in connection with the total wanted fertility rate (WTFR) indicator. The alternative definitions of wanted births noted in connection with the WTFR may also be applied to the present indicator.

Chapter X

Future Steps

- Inclusion of impact measures other than fertility
- Refinement of indicators in functional areas
- Inclusion of cost indicators
- Inclusion of indicators for programs directed to special target audiences
- Development of scoring rules for indicators
- Inclusion of indicators of institutionalization and sustainability

Chapter X

FUTURE STEPS IN THE DEVELOPMENT OF INDICATORS FOR FAMILY PLANNING EVALUATION

This Handbook provides an inventory of indicators to be used in evaluating family planning programs in terms of process and output (using program–based data) and outcomes (using population–based data). In this first edition of the Handbook, we have included those indicators that are currently in use or for which there is some degree of consensus. Topics to be further developed in the second edition, scheduled for 1995, include the following.

Inclusion of Impact Measures other than Fertility

The current edition of the Handbook focuses on fertility as the ultimate outcome expected from family planning programs. In fact, this motive has been central to the promotion of family planning in many developing countries around the world.

However, fertility is not the only reason for promoting family planning. There is a growing body of empirical evidence that documents the positive effect of: (a) increased intervals between births, (b) avoidance of births under 20 and over 35 years of age, and (c) the avoidance of high parity births, on both maternal and child health (National Research Council, 1989). Moreover, the ability to control one's fertility is viewed as a basic human right, which is both a determinant and consequence of the status of women in a given country. Donor agencies, policy makers, and program managers are now focusing increased attention on outcomes related to health, women's status, and reproductive rights, topics to be covered in the second edition of the Handbook.

Refinement of Indicators in Functional Areas

As mentioned in Chapter I, a series of Working Groups have been organized under The EVALUATION Project to develop and test lists of indicators in the functional areas outlined in this manual. Their work is reflected in Chapters II – IV of this Handbook.

However, much of this work is still in progress. The meetings of these groups have been staggered, such that some groups are further along with the task than others (a fact that reflects on the timing of the meetings, not the productivity of the groups).

Only two working groups (Training, and Commodities and Logistics) have completed the first stage of this process: the development of a list of indicators for the functional area. Yet testing of these indicators at the field level remains to be done. The Service Delivery Working Group has produced the indicators for Quality of Care (see Appendix C), which are also at the field testing stage; the group is now focusing on the measurement of cost, accessibility, and service utilization. The Operations Research Working Group has developed a first draft of indicators, but has yet to meet to refine these or to establish a system for testing them.

Two other groups (Management and Policy) are in progress as this edition of the Handbook goes to press. Thus, the sections on these topics in this Handbook constitute drafts to be discussed and reworked with these groups, rather than a definitive set of indicators. Finally, the working groups on Information–Education–Communication and on Evaluation have yet to be formed. The sections on these topics have been developed not through the group process but rather with input from key individuals in this area.

In short, the work in further developing and in field testing the indicators for the different functional areas will no doubt result in modifications to the measures listed in this first edition. Moreover, we anticipate further experimentation and testing of composite indicators, such as the one described

for Commodities and Logistics. These refinements will be incorporated into the second edition.

Inclusion of Cost Indicators

In this edition, indicators of cost are conspicuously lacking. However, the international family planning community is growing increasingly concerned with issues of efficiency, sustainability, and cost recovery, among others. The second edition will address these topics and provide indicators for use in cost analyses.

Inclusion of Indicators for Programs Directed to Special Target Audiences

Although a number of the indicators in the Handbook are applicable to programs directed toward specific subgroups (for example, adolescents), this first edition mirrors the trend in international family planning programs to consider married couples (or women) as the target audience. Thus, there is a strong emphasis on utilization of family planning services and contraceptive practice. The second edition will broaden this focus to consider the characteristics of programs directed to special populations that would in turn affect the choice of evaluation indicators.

Development of Scoring Rules for Indicators

As mentioned in Chapter I, the current edition does not propose rules for quantifying specific indicators or sets of indicators, which are potentially useful with respect to the components of the family planning supply environment (e.g., for policy environment, service delivery operations, and service outputs). This issue will be addressed in the next edition, based on further indicator testing at the field level.

Inclusion of Indicators of Institutionalization and Sustainability

Family planning programs worldwide run the gamut from "emergent" to "mature" (Destler et al., 1990). In those countries with fledgling family planning programs, much of the USAID assistance has gone to designing and implementing family planning programs. By contrast, as programs mature, increasing importance is placed on institutionalization and sustainability.

A key question is then how to evaluate the progress of a program with respect to institutionalization and sustainability. Although these concepts are discussed extensively in the literature, there has been relatively little attempt to measure these dimensions in the context of international family planning programs. The EVALUATION Project in conjunction with Family Health International organized a meeting to look at institutionalization of training in family planning programs, the minutes of which are currently available (The EVALUATION Project, 1992c). However, much work remains to be done on this topic.

The second edition of the Handbook will examine this issue and propose indicators to measure institutionalization and sustainability in international family planning programs.

References and Appendices

References

A.I.D. Task Force on Standardization of Family Planning Program Performance Indicators, "Summary of Task Force Discussions and Recommendations," January 1987, draft report.

A.I.D. Task Force on Standardization of Family Planning Program Performance Indicators, "Revised CYP factors," August 1991, draft report.

Anderson, B. and B. Silver. 1992. "A Simple Measure of Fertility Control," *Demography* 29, 3:343–356.

Arnold F. and A. Blanc. 1989. *Fertility Levels and Trends*. Demographic and Health Surveys Comparative Studies No. 2., Columbia, MD: Institute for Resource Development/Macro International, Inc.

Baldwin, G. 1992. *Targets and Indicators in World Bank Population Projects*. Policy Research Working Papers Population, Health, and Nutrition. The World Bank, Population and Human Resources Department.

Bertrand, J.T. and L.F. Brown. 1993. *Final Report: Working Group on the Evaluation of Family Planning Training.* The EVALUATION Project, UNC–CH: Carolina Population Center.

Blanc, A.K. and N. Rutenberg. 1990. "Coitus and Contraception: The Utility of Data on Sexual Intercourse for Family Planning Programs, " *Studies in Family Planning* 22, 3:162–176.

Blumenfeld, S.N. 1985. *Operations Research Methods: A General Approach to Primary Health Care*. Chevy Chase, Maryland: PRICOR.

Bongaarts, J. 1978. "A Framework for Analyzing the Proximate Determinants of Fertility", *Population and Development Review* 4, 2:105-132.

Bongaarts, J. and J. Menken. 1983. "The Supply of Children: A Critical Essay." In *Determinants of Fertility in Developing Countries*, Vol. 1:1–26. New York: Academic Press.

Bongaarts, J. 1984. "A Simple Method for Estimating the Contraceptive Prevalence Required to Reach a Fertility Target." *Studies in Family Planning* 15, 4:184–190.

Bongaarts, J. 1985. "The Concept of Potential Fertility in the Evaluation of the Fertility Impact of Family Planning Programs." In *Studies to Enhance the Evaluation of Family Planning Programmes*, 40–49. New York: Department of International Economic and Social Affairs.

Bongaarts, J. 1990. "The Measurement of Wanted Fertility," *Population and Development Review* 16, 3:487–506.

Bongaarts, J. and G. Rodriguez 1991. "A New Method for Estimating Contraceptive Failure Rates." In United Nations *Measuring the Dynamics of Contraceptive Use*. New York: United Nations, Department of International Economic and Social Affairs.

Bongaarts, J. 1991a. "Do Reproductive Intentions Matters?" In *Proceedings of the Demographic and Health Surveys World Conference*. Columbia, Maryland: IRD/Macro International.

Bongaarts, J. 1991b. "The KAP–gap and the Unmet Need for Contraception." *Population and Development Review* 17, 2: 293–313.

Boulier, B. 1985. The Effects of Family Planning Programs on Fertility in the Developing World. World Bank Staff Papers, No. 677.

Bruce, J. 1990. "Fundamental Elements of Quality of Care: A Simple Framework." *Studies in Family Planning* 21, 2:61–90.

Bulatao, R.A. 1981. "Values and Disvalues of Children in Successive Childbearing Decisions," *Demography* 18: 1–25.

Bulatao, R.A. and R.D. Lee. 1983. *Determinants of Fertility in Developing Countries*. New York: Academic Press.

CEDPA and FHI. 1992. *Service Quality Assessment and Improvement Processes and Tools September 23–24*. Washington, DC

Centers for Disease Control. 1985. *Logistics Guidelines for Family Planning Programs*, Atlanta: CDC Division of Reproductive Health.

Chandrasekaran, C. and A.I. Hermalin (eds). 1975. *Measuring the Effect of Family Planning Programs on Fertility*. Dolhain, Belgium: Ordina Editions.

Chayovan, N., A. I. Hermalin and J. Knodel. 1984. "Measuring Accessibility to Family Planning Services in Thailand." *Studies in Family Planning* 15, 5:201–211.

Coeytaux, F.M., T. Kilani, and M. McEvoy. 1987. "The Role of Information, Education, and Communication in Family Planning Service Delivery in Tunisia." *Studies in Family Planning* 18, 4:229-233.

Crane, B. and J. Finkle. 1990. "The Politics of International Population Policy." In *International Transmission of Population Policy Experience*, 167–183. New York: United Nations Department of International Economic and Social Affairs.

Demeny, P. 1988. "Social Science and Population Policy," *Population and Development Review* 14, 3:451–470.

Destler, H., D. Liberi, J. Smith, and J. Stover. 1990. *Preparing for the Twenty–First Century: Principles for Family Planning Service Delivery in the Nineties*. Washington DC: Family Planning Services Division Office of Population U.S. Agency for International Development.

Dixon–Mueller, R. and A. Germain. 1992. "Stalking the Elusive 'Unmet Need' for Family Planning." *Studies in Family Planning* 23, 5:330–335.

Easterlin, R.A. 1978. "The Economics and Sociology of Fertility: A Synthesis." In *Historical Studies of Changing Fertility*, Princeton: Princeton University Press.

Elkins, H. and O.S. Nordberg. 1977. "Service Statistics: Aid to more Effective Family Planning Program Management." *Population Reports*, Series J. No. 17. George Washington University Medical Center.

Entwisle, B. 1989. "Measuring the Components of Family Planning Program Effort." *Demography* 269, 1:53–80.

Espenshade, T. 1977. "The Taste for Children." *Family Planning Perspectives* 9:40–47.

The EVALUATION Project. 1992a. *Report of the First Meeting of the Commodities and Logistics Working Group*. Chapel Hill, North Carolina: Carolina Population Center.

The EVALUATION Project. 1992b. "Minutes of the Meeting on Service Delivery Working Group, June 3-4, 1992," and "Minutes of the Meeting Subcommittee on Quality Assurance, Service Delivery Working Group, June 17, 1992." Chapel Hill, North Carolina: Carolina Population Center.

The EVALUATION Project. 1992c. "Minutes of the Meeting on Institutionalization of Training, October 28, 1992." Chapel HIll, North Carolina: Carolina Population Center.

The EVALUATION Project. 1992d. "Minutes of the Meeting of the Working Group on Operations Research, November 12, 1992." Chapel Hill, North Carolina: Carolina Population Center.

The EVALUATION Project. 1993a. *Report of the Second Meeting of the Commodities and Logistics Working Group*. Chapel Hill, North Carolina: Carolina Population Center.

The EVALUATION Project. 1993b. "Minutes of the Meeting of the Management Working Group, June 14, 1993." Chapel Hill, North Carolina: Carolina Population Center.

Finkle, J. and G. Ness. 1985. *Managing Delivery Systems: Identifying Leverage Points for Improving Family Planning Program Performance*. Final Report, Managing Delivery Systems Project, AID Contract No. DPE–0632–C–00–3065–00.

Fisher, A.A., J.E. Laing, J.E. Stoeckel and J.W. Townsend. 1991. *Handbook for Family Planning Operations Research Design*. New York: The Population Council.

Fisher, A., B. Mensch, R. Miller, S. Askew, A. Jain, C. Ndeti, L. Ndhlovu, and P. Tapsoba. 1992. *Guidelines and Instruments for a Family Planning Situation Analysis Study*. New York: The Population Council.

Foreit, J.R., M.E. Gorosh, D.G. Gillespie, and G.C. Merritt. 1978. "Community–based and Commercial Contraceptive Distribution: An Inventory and Appraisal." *Population Reports*, Series J. No. 19. March:J1–J29.

Foreit, K. 1992. "Unmet Demand for Contraception vs. Unmet Demand for Appropriate Contraception." Paper Presented at the 120th Annual Meeting of the American Public Health Association, Washington, DC, November.

García–Nuñez, J. 1992. Improving Family Planning Evaluation: A Step–by–step Guide for Managers and Evaluators. West Hartford, Connecticut: Kumarian Press, Inc.

Gaslonde, S. and E. Carrasco. 1982. "The Impact of Some Intermediate Variables on Fertility: Evidence from the Venezuela National Fertility Survey 1977." *WFS Occasional Papers*, No. 23. London: World Fertility Survey.

Hauser, P.M. 1967. "Family Planning Programs and Population Programs: A Book Review Article." *Demography* 4, 1:397–414.

Hermalin, A.I. 1983. "Fertility Regulation and its Costs: A Critical Essay." In *Determinants of Fertility in Developing Countries*, Vol. II, 1–53. New York: Academic Press.

Hermalin, A.I. and B. Entwisle. 1985. "Future Directions in the Analysis of Contraceptive Availability." Proceedings from the International Population Conference, Florence, 5–12 June 1985.

Hermalin, A.I. and B. Entwisle. 1988. "The Availability and Accessibility of Contraceptive Services." In R.J. Lapham and W.P. Mauldin *Organizing for Effective Family Planning Programs*. Washington, DC: National Academy Press.

Hermalin, A. I., A.P. Riley and L. Rosero–Bixby. 1992. "A Multi–level Analysis of Contraceptive Use and Method Choice in Costa Rica." Population Studies Center, University of Michigan.[unpublished]

Hofstede, G. 1982. "National Cultures in Four Dimensions." *International Studies of Management and Organization*, XII 1–2: 46–74.

Hutchings, J., G. Perkin, and L. Saunders. 1987. "The Effect of Contraceptive Technology on the Program Environment." In R.J. Lapham and G.B. Simmons (ed.) Organizing for Effective Family Planning Programs. Washington, DC: National Academy Press.

ICDDR, B. 1989. "Why Contraceptive Prevalence Estimates Based on Service Statistics and Sample Surveys Do Not Agree." MCH–FP Extension Project Briefing Paper, No. 11.

Ickis, J. 1987. "Structural Issues Related to Delivery Systems." In R.J. Lapham and G. Simmons (eds) *Organizing for Effective Family Planning Programs*. Washington, DC: National Academy Press.

Jain, A. 1989. "Fertility Reduction and the Quality of Family Planning Service." *Studies in Family Planning* 20, 1:1–16.

Jain, A (ed.), 1992. *Managing Quality of Care in Population Programs*. West Hartford: Kumarian Press.

Jain, A, J. Bruce, and B. Mensch. 1993. "Setting Standards of Quality in Family Planning Programs." *Studies in Family Planning* 23, 6:392–395.

Jejeebhoy, S. 1989. "Measuring the Quality and Duration of Contraceptive Use: An Overview of New Approaches." *Population Bulletin of the United Nations*, No. 26.

Jemai, Y. 1985. "Application of the prevalence model: the case of Tunisia." In *Studies to Enhance the Evaluation of Family Planning Programmes*, 234–236. New York: United Nations Department of International Economic and Social Affairs.

Kendall, Sir Maurice. 1979. "The World Fertility Survey: Current Status." *Population Reports*, Series M. No.3.

Kincaid, D.L., P.T. Piotrow, J.G. Rimon, B. Lozare. 1993. "Conceptual Frameworks for the Strategic Design of Health Communication Projects." Working Paper, Center for Communication Programs, Department of Population Dynamics, The Johns Hopkins University.

Kincaid, D.L. Personal Communication. August 1992.

Knodel, J. 1983. "Natural Fertility: Age Patterns, Levels and Trends." In *Determinants of Fertility in Developing Countries*, Vol. I 26–60. New York: Academic Press.

Knowles, J.C., K.A. Bollen and K. Yount. 1993. "The Policy Environment of Family Planning Programs: A Literature Review." The EVALUATION Project. Chapel Hill North Carolina: Carolina Population Center.

Labbok, M. 1992. "The Lactational Amenorrhea Method: Why Bother? Clinical Data for Policy Change," presented at the Annual Meeting of the American Public Health Association, Washington, DC, Nov. 8–12, 1992.

Laing, J. 1985. "Continuation and Effectiveness of Contraceptive Practice: A Cross–sectional Approach." *Studies in Family Planning* 16, 3:138–153.

Landry, E. (Chair, Evaluation Working Group). "Final Report Outcome of the Evaluation Group." Prepared for the Information and Training Steering Committee, USAID, February 1992.

Lapham, R.J. and W.P. Mauldin. 1984. "Family planning program effort and birthrate decline in developing countries." *International Family Planning Perspectives* 10, 4:109–118.

Lapham, R. J. and W.P.Mauldin. 1985. "Contraceptive Prevalence: The Influence of Organized Family Planning Programs." *Studies in Family Planning* 16, 3:117–137.

Lapham, R. and G. Simmons. 1987. "Overview and Framework." In *Organizing for Family Planning Program Effectiveness*. Washington, DC: National Academy Press.

Laurent, A. 1983. "The Cultural Diversity of Western Conceptions of Management." *International Studies of Management and Organization*, XIII 1–2: 75–96.

Lee, R.D. 1980. "Aiming at a Moving Target: Period Fertility and Changing Reproductive Goals." *Population Studies* 34: 205–226.

Lightbourne, R.E. 1985. "Desired Number of Births and Prospects for Fertility Decline in 40 Countries." *International Family Planning Perspectives* 11, 2:34–47.

Lightbourne, R.E. 1987. "Reproductive Preferences and Behavior." In J. Cleland and C. Scott (eds) *The World Fertility Survey: An Assessment*. London: Oxford University Press.

London, K.A., J. Cushing, S.O. Rutstein, et al. 1985. "Fertility and Family Planning Surveys: An Update." *Population Reports*, Series M. No.8.

Lutz. W. 1990. *Distributional Aspects of Human Fertility: A Global Comparative Study*. New York: Academic Press.

Mauldin, W.P. and J.A. Ross. 1991. "Family Planning Programs: Efforts and Results, 1982–89." *Studies in Family Planning* 22, 6:350–367.

McClelland, G.H. 1983. "Family Size Desires as Measures of Demand." In *Determinants of Fertility in Developing Countries*. New York: Academic Press.

McGuire, W. J. 1989. "Theoretical Foundations of Campaigns." In *Public Communication Campaigns*, Second Edition. Newbury Park: Sage Publications.

Mensch, B. and A. Jain. 1991. "Measuring Quality of Care at the Policy Level." Paper Presented at the American Public Health Association Meeting, Atlanta, Georgia, November 1991.

Mensch, B., A. Fisher, I. Askew, and A. Ajayi. 1993. "Supply Side Determinants of Family Planning Clinic Utilization: Findings from African Situation Analysis Studies." Paper Presented at the Population Association of America Meetings, Cincinnati, Ohio, April 1993.

Morris, L. and J.E. Anderson. 1982. "The Use of Contraceptive Prevalence Survey Data to Evaluate Family Planning Program Service Statistics." In *The Role of Surveys in the Analysis of Family Planning Programs*. Liege, Belgium: Ordina Editions.

Morris, L., G. Lewis, D.L. Powell, J. Anderson, A. Way, J. Cushing, and G. Lawless. 1981. "Contraceptive Prevalence Surveys: A New Source of Family Planning Data." *Population Reports*, Series M. No.5.

Namboodiri, N.K. and C.M. Suchindram. 1987. *Life Table Techniques and Their Applications*. New York, NY: Academic Press.

National Research Council. 1989. *Contraception and Reproduction. Health Consequences for Women and Children in the Developing World*. Washington, DC: National Academy Press.

Pearl, R. 1932. "Contraception and Fertility in 2000 Women." *Human Biology*, 4:363–407.

Piotrow, P.T., J.G. Rimon II, K. Winnard, D.L. Kincaid, D. Huntington, and J. Convisser. 1990. "Mass Media Family Planning Promotion in Three Nigerian Cities." *Studies in Family Planning* 21, 5:265-274.

Piotrow, P.T., D.L. Kincaid, M. J. Hindin, C.L. Lettenmaier, I. Kuseka, T. Silberman, A. Zinanga, F. Chikara, D.J. Adamchak, M.T. Mbizvo, W. Lynn, O.M. Kumah, and Y. Kim. 1992. "Changing Men's Attitudes and Behavior: The Zimbabwe Male Motivation Project." *Studies in Family Planning* 23, 6:365–375.

Prasartkul, P., Y. Porapakkham and W. Sittitrai. 1987. *Report on Birth Order Distribution as a Family Planning Program Evaluation Indicator*. Bangkok, Thailand: Mahidol University, Institute for Population and Social Research.

Pullum, T.W. 1980. *Illustrative Analysis: Fertility Preferences in Sri Lanka*. WFS Scientific Reports, No. 9. London: World Fertility Survey.

"Report of the Subcommittee on Quality Indicators in Family Planning Service Delivery." Prepared for AID's Task Force on Standardization of Family Planning Program Performance Indicators. October 1990.

Reynolds, J. 1990. "Evaluation of Child Survival Programs." In *Health Care of Women and Children in Developing Countries*, 154–164. Oakland, California: Third Party Publishing.

Robey, B., S.O. Rutstein, L. Morris, and R. Blackburn. 1992. "The Reproductive Revolution: New Survey Findings." *Population Reports*, Series M. No.11.

Rogers, E.M. 1973. *Communication Strategies for Family Planning*, New York: The Free Press.

Ross, J. (ed.) 1992. *International Encyclopedia of Population.* Volume 1. New York: The Free Press.

Ross, J.A., M. Rich, J.P. Molzan and M.l Pensak. 1988. *Family Planning and Child Survival: 100 Developing Countries*. New York: Columbia University Center for Population and Family Health.

Ross, J.A., W.P. Mauldin, Steven R. Green and E. Romana Cooke. 1992. *Family Planning and Child Survival Programs, as Assessed in 1991*. New York: The Population Council.

Ryder, N. 1982. *Progressive Fertility Analysis*. World Fertility Survey Technical Bulletin No. 8. Voorburg, Netherlands: International Statistical Institute.

Shelton, James. 1991. "What's Wrong with CYP?" *Studies in Family Planning* 22, 5:332-335.

Shelton, J.D., M.A. Angle, and R.A. Jacobstein. 1992. "Medical Barriers to Access to Family Planning." *Lancet*. 340:1334–5.

Sherris, Jacqueline D, Kathy A. London, Sidney H. Moore, John H. Pile and Walter B. Watson, 1985. "The Impact of Family Planning Programs on Fertility." *Population Reports*, XIII, 1:J733-J771.

Srinivasan, K. and M. Freymann. 1989. "Need for a Reorientation of Family Planning Program Strategies in Developing Countries: A Case for a Birth-based Approach." In *Dynamics of Population and Family Welfare*. Bombay, India: Himilaya Publishing House.

Srinivasan, K. P.C. Saxena and A. Pandey. 1992. "Birth Order and Birth Interval Statistics." *International Institute for Population Sciences*. Bombay, India.[unpublished]

Stover, J., J.T. Bertrand, S. Smith, and N. Rutenberg. 1993. "Empirically–based Conversion Factors for Calculating Couple–Years of Protection." Chapel Hill, N.C.: The EVALUATION Project (unpublished).

Stover, J., J.C. Knowles, A.E. Wagmen, D.L. Nortman, J. Freyman, and R. McKinnon. 1991. *Target Cost: A Model for Projecting the Family Planning Service Requirements and Costs to Achieve Demographic Goals*. Washington, DC: The Futures Group.

Subcommittee on Quality Indicators in Family Planning Service Delivery. 1990. Report of the Subcommittee. Submitted to the Agency for International Development's Task Force on Standardization of Family Planning Program Performance Indicators.

Tsui, Amy O. and Pamina Gorbach. 1993. *Conceptual Framework for the Evaluation of Family Planning Programs*. Chapel Hill, North Carolina: Carolina Population Center.

Tsui, A.O., A.I. Hermalin, J.T. Bertrand, J. Knowles, J. Stover, and K. Stewart. 1992. "Evaluating Family Planning Program Impact: New Initiatives on a Persisting Question." Paper presented at the Annual Meeting of the Population Association of America, April 30–May 2, 1992, Denver, Colorado.

Tsui, A.O., and L.H. Ochoa. 1992. "Service Proximity as a Determinant of Contraceptive Behavior: Evidence from Cross–national Studies of Survey Data." In J.F. Philipps and J.A. Ross (eds) *Family Planning Programs and Fertility*. London: Clarendon Press–Oxford.

United Nations. 1979. *Manual IX, The Methodology of Measuring the Impact of Family Planning Programmes on Fertility*. New York: Department of International Economic and Social Affairs.

United Nations. 1982. *Evaluation of the Impact of Family Planning Programmes on Fertility: Sources of Variance*. New York: Department of International Economic and Social Affairs.

United Nations. 1985. *Studies to Enhance the Evaluation of Family Planning Programmes*. New York: Department of International Social and Economic Affairs.

United Nations. 1986. *Manual IX, The Methodology of Measuring the Impact of Family Planning Programmes on Fertility – Addendum*. New York: Department of International Social and Economic Affairs.

United Nations. 1991. *The Dynamics of Contraceptive Use*. New York: Department of International Social and Economic Affairs.

Veney, J. 1992. *Definition of Terms: Types of Evaluation*. Chapel Hill, North Carolina: Carolina Population Center, The EVALUATION Project.

Ward, V.M., J.T. Bertrand, and F. Puac. 1992. "Exploring Socialcultural Barriers to Family Planning among Mayans in Guatemala. *International Family Planning Perspectives*. 18(2):59–65.

Warwick, D. 1988. "Culture and the Management of Family Planning Programs." *Studies of Family Planning Programs* 19, 1:1–18.

Weller, R.H., D. F. Sly, A. Sukamdi, and R. Ekawati. 1981. "The Wantedness Status of Births in Indonesia." In *Proceedings of the Demographic and Health Surveys World Conference*. Columbia, Maryland: IRD/Macro International.

Westoff, C. 1990. "Reproductive Intentions and Fertility Rates." *International Family Planning Perspectives* 16, 3:84-96.

Westoff, C.F. 1991. *Reproductive Preferences*. Demographic and Health Surveys Comparative Studies No. 3. Columbia, Maryland: Institute for Resource Development/Macro International, Inc.

Westoff, C.F. and L.H. Ochoa. 1991. *Unmet Need and Demand for Family Planning*. Demographic and Health Surveys Comparative Studies No. 5., Columbia, Maryland: Institute for Resource Development/Macro International, Inc.

Wishik, S.M. and K. Chen. 1973. "Couple–years of Protection: A Measure of Family Planning Program Output." New York: Columbia University, International Institute for the Study of Human Reproduction.

Listing of All Indicators by Primary Source of Data

Source: Administrative and Program Records

Policy Environment

- Existence of a policy development plan
- Number of appropriately disseminated policy analyses
- Number of awareness–raising events targeted to leaders
- Existence of a strategic plan for expanding the national family planning program
- Integration of demographic data into development planning
- Number of statements of leaders in support of family planning
- Formal population policy addressing fertility and family planning
- National family planning coordination
- Level of the family planning program within the government administration
- Levels of import duties and other taxes
- Restrictions on advertising of contraceptives in the mass media
- Absence of unwarranted restrictions on providers
- Absence of unwarranted restrictions on users
- Public sector resources devoted to family planning as a percentage of GDP

Service Delivery Operations (functional areas)

Management

- Existence of a clear mission that contributes to the achievement of program goals
- Realization of operational targets
- Clearly defined organizational structure
- Adequacy of staffing
- Awareness of current financial position
- Access to current information on key areas of program functioning
- Access to current information on program progress
- Capacity to track commodities

Training

- Number/percentage of courses that achieve learning objectives
- Number/percentage of courses that contribute to the achievement of program training objectives
- Number of trainees by type
- Number/percentage of trainees who apply the skills to their subsequent work

SOURCE: ADMINISTRATIVE AND PROGRAM RECORDS, CONTINUED

Commodities and Logistics

- Pipeline wastage
- Frequency of stock–outs
- Percentage of key personnel trained in contraceptive logistics
- Composite indicator for commodities and logistics

Information – Education – Communication

- Number of communications produced, by type, in a reference period
- Number of communications disseminated, by type, during a reference period

Research and Evaluation

- Presence of an active research and evaluation unit
- Extent of use of a service statistics system
- Conduct of periodic household and/or special purpose surveys and studies
- Conduct of operations research (OR)
- Regular conduct of process evaluations
- Conduct of effectiveness, efficiency, and impact evaluations
- Use of research and evaluation results for program modification
- Dissemination of research and evaluation results

Service Outputs

Accessibility (illustrative)

- Cost of one month's supply of contraceptives as a percentage of monthly wages
- Restrictive program policies on contraceptive choice

Program Image

- Number and type of activities to improve the public image of family planning during a reference period (e.g., one year)

SOURCE: SURVEY DATA (E.G., DHS)

Policy Environment

- Extent of commercial sector participation[1]

---

[1] A new interactive software package, EASEVAL, has been developed under The EVALUATION Project in conjunction with the DHS staff, to facilitate the use of the DHS data from standard recode files. The package is currently available for use with DHS I files; it will be available for use with DHS II once the standard recode files for DHS II are released. The menu of this software allows the user to select "INDICATORS," which gives a complete listing of the indicators described in this Handbook that are available from the DHS survey. In Appendix A, all indicators available from the DHS interactive software package are marked with a superscript[1] to highlight the link between this Handbook of Indicators and the EASEVAL package.

SOURCE: SURVEY DATA (E.G., DHS), CONTINUED

Demand for Children

- Mean desired family size[1]
- Desire for additional children[1]
- Wanted status of previous births[1]
- Wanted total fertility rate (WTFR)[1]

Demand for Family Planning

- Demand for limiting
- Demand for spacing
- Total demand (for family planning)[1]
- Unmet need for family planning[1]
- Satisfaction of demand for family planning

Service Outputs

Accessibility

- Percentage of the target population who know at least one source of contraceptive services/supplies
- Percentage of non–use related to psycho–social barriers

Program Image

- Percentage of the target population favorable to the (national) family planning program

Contraceptive Practice

- Contraceptive prevalence rate (CPR)[1]
- Number of current users
- Level of ever (past) use[1]
- Source of supply (by method)[1]
- Method mix[1]
- User characteristics
- Continuation rates
- Use failure rates

Fertility

Fertility Level

- Crude birth rate (CBR)
- Age specific fertility rate (ASFR)[1]
- Total fertility rate (TFR)[1]

SOURCE: SURVEY DATA (E.G., DHS), CONTINUED

Births Averted

- Births averted (by the program)

Other Indicators

- Parity–specific birth rate[1]
- Proportion of births above (or below) a specified parity[1]
- Proportion of births by women above or below a specified age[1]
- Median length of birth intervals[1]
- Proportion of open or closed birth intervals that are of a specified length or longer[1]
- Unwanted total fertility rate (UTFR)[1]

SOURCE: SERVICE STATISTICS, MANAGEMENT INFORMATION SYSTEMS

Service Delivery Operations (functional areas)

Management

- Management awareness of current financial position
- Access to current information on key areas of program functioning
- Capacity to track commodities

Commodities/Logistics

- Pipeline wastage
- Percentage of storage capacity meeting acceptable standards
- Frequency of stock–outs
- Percent of service delivery points (SDPs) stocked according to plan
- Composite indicator for commodities and logistics

Service Utilization

- Number of visits to service delivery points
- Number of acceptors new to modern contraception
- Number of acceptors new to the institution
- Number of new segment acceptors
- Couple–years of protection (CYP)
- Method mix
- User characteristics
- Continuation rates

SOURCE: SPECIAL SURVEYS (OF PROVIDERS, CLIENTS, TARGETED AUDIENCES, ETC.)

Policy Environment

- Extent of commercial sector participation

Service Delivery Operations (functional areas)

Management

- Clearly defined organizational structure
- Adequacy of staffing
- Awareness of current financial position
- Access to current information on key areas of program functioning
- Access to current information on program progress

Training

- Number of trainees who have mastered relevant knowledge
- Number/percentage of trainees who apply the skills to their subsequent work

Information – Education – Communication

- Percentage of target audience exposed to program messages, based on respondent recall
- Percentage of target audience who correctly comprehend a given message
- Percent of audience who acquire skill to complete a certain task as a result of exposure to a specific communication
- Percentage of target audience exposed to a specific message who report liking it
- Number/percentage of target audience who discuss message(s) with others, by type of person
- Percentage of target audience who advocate family planning practice

Research

- Extent of use of service statistics system
- Use of research and evaluation results for program modification

Service Outputs

Accessibility

- Number of SDPs located within a fixed distance or travel time of a given location (i.e., service density)
- Cost of one month's supply of contraceptives as a percentage of monthly wages

[2] This category also includes interviews with managers and other staff, responses of participants on training evaluations, and interviews with key informants.

SOURCE: SPECIAL SURVEYS (OF PROVIDERS, CLIENTS, TARGETED AUDIENCES, ETC.), CONTINUED

Quality

- Percentage of clients reporting "sufficient time" with provider
- Percentage of clients who perceive that hours/days are convenient
- Percentage of clients informed of timing and sources for resupply/revisit

Program Image

- Percentage of target population favorable to the (national) family planning program

SOURCE: OBSERVATION/ASSESSMENT BY EXPERTS

Policy Environment

- Quality of program leadership

Service Delivery Operations (functional areas)

Training

- Number/percentage of courses that achieve learning objectives
- Number/percentage of courses where training methodology is appropriate for the transfer of skills and knowledge
- Number/percentage of trainees who have mastered relevant knowledge
- Number/percentage of trainees competent to provide a specific family planning service
- Number/percentage of trained providers assessed to be competent at a specific period (e.g., six months) post–training

Commodities/Logistics

- Composite indicator for commodities and logistics

Research and Evaluation

- Extent of use of a service statistics system

Service Outputs

Quality

- Number of contraceptive methods available at a specific SDP
- Percentage of counseling sessions with new acceptors in which provider discusses all methods
- Percentage of client visits during which provider demonstrates skill at clinical procedures, including asepsis
- Percentage of clients informed of timing and sources for resupply/revisit

INDICATORS FOR EVALUATING THE MANAGEMENT FUNCTIONAL AREA[1]

ELEMENT/INDICATOR

Element No. 1 – Program Mission

1.1 Existence of a clear program mission that contributes to the achievement of national family planning goals.

1.2 The mission is understood throughout the program/organization.

1.3 The program mission is used in planning and operations by each level/segment of the organization.

Element No. 2 – Planning/Strategy

2.1 Strategic/operational plans are realized

2.2 Existence of a strategic plan

2.3 The strategic plan is understood and used

2.4 Periodic review of the strategic plan

2.5 Existence of operational plans

2.6 Operational plans are used

2.7 Operational plans are regularly reviewed and updated

Element No. 3 – Organization

Sub–Element 3.1 Organizational structure

- 3.1.1 Clear organizational structure
- 3.1.2 Defined decision–making process
- 3.1.3 System for communicating decisions
- 3.1.4 Formal reporting networks
- 3.1.5 Written position descriptions
- 3.1.6 Existence of written service standards/guidelines
- 3.1.7 Management subsystems are integrated, coherent, and consistent with each other
- 3.1.8 Program is able to regenerate leadership
- 3.1.9 Program is able to survive change in leadership

Sub–Element 3.2 Leadership

- 3.2.1 Program mission and strategic plan are forward looking
- 3.2.2 Leadership is able to adapt to changes in the external environment

[1] The EVALUATION Project, 1993b

	3.2.3	Leadership is able to influence the external environment
	3.2.4	Internal influence
	3.2.5	Leadership is able to communicate mission/strategy
Sub–Element	3.3	Decision–making
	3.3.1	Information–based decision making
	3.3.2	Decisions are made and communicated in timely manner
	3.3.3	Decision–makers solicit input internally/externally (information seeking)

Element No. 4 – Human Resources

Sub–Element	4.1	Clearly defined position descriptions
	4.1.1	Written position descriptions
Sub–Element	4.2	Interaction/communication guidelines exist for decision making/implementation
	4.2.1	Mechanisms for problem solving
	4.2.2	Mechanisms that facilitate feedback
	4.2.3	Mechanisms for performance review
Sub–Element	4.3	Recruitment/staffing
	4.3.1	Adequacy of staffing
	4.3.2	Staffing decisions are technically based and timely
	4.3.3	Staff have appropriate qualifications to execute responsibilities
Sub–Element	4.4	Training
	4.4.1	Rational use of training
	4.4.2	Fairness in use of training opportunities
Sub–Element	4.5	Career Orientation
	4.5.1	Clear definition of professional growth tracks
Sub–Element	4.6	Compensation policies
	4.6.1	Policies are merit–based
	4.6.2	Compensation levels are competitive within environment
	4.6.3	Use of performance–based incentives
	4.6.4	Fairness

Element No. 5 – Finance

Sub–Element	5.1	Financial management systems
	5.1.1	Financial management procedures
	5.1.2	Information system
Sub–Element	5.2	Revenue generation

5.2.1 Adequacy of revenues

5.2.2 Revenue generation mechanisms

5.2.3 Ability to raise projected revenues

5.2.4 Diversity of funding

Element No. 6 – Information

Sub–Element 6.1 Adequacy of information support

6.1.1 Access to current information on key areas of program functioning

6.1.2 Relevant information systems (sub–systems) are in place

6.1.3 Relevance of information collected

6.1.4 Reporting burden

Sub–Element 6.2 Use of Information for Management Decisions

6.2.1 Information system is based upon the identification of users and opportunities for use of information

Sub–Element 6.3 Timeliness

6.3.1 Information is produced as needed

6.3.2 Extent to which information respects the schedule

Sub–Element 6.4 Quality of information (externally defined)

6.4.1 Sensitivity of data

6.4.2 Reliability/accuracy of data

Element No. 7 – Monitoring and Evaluation

7.1 Availability of current information on progress made toward the accomplishment of program targets, objectives, and goals

7.2 Establishment of a monitoring and evaluation plan

7.3 Establishment of a monitoring system

7.4 Timeliness of monitoring/evaluation systems

7.5 Use of monitoring and evaluation data

7.6 Evaluation goals/objectives achieved

Element No. 8 – Logistics

8.1 Capacity to track commodities

8.2 Existence of a logistics system

8.3 Existence of an LMIS

8.4 System reviewed periodically at the top level

8.5 Someone at each level is responsible for logistics management

8.6 Plans for logistics system maintenance are in place

Quality Indicators for Clinic–Based Family Planning Programs

Subcommittee on Quality
Service Delivery Working Group (SDWG)
The EVALUATION Project

Notes on the Indicators

The following list of indicators was initially drafted by the Quality Subcommittee on June 17, 1992. Since then, it has undergone a series of revisions that are reflected in this version. In addition, analagous lists of quality indicators have been developed for community based distribution and social marketing projects (The EVALUATION Project, 1992c).

At the June 1992 meeting the group concurred that there are three levels for measuring indicators related to quality: (1) manager (referred to elsewhere as the "sub–systems" or "enabling systems"), (2) provider, and (3) client. These levels are interrelated, since certain inputs must be in place at the manager level (e.g., a full range of contraceptive methods in stock) to enable the provider to take the correct action (e.g., offer the client all methods that are medically appropriate), which can in turn be measured at the client level (e.g., by asking the client what methods were offered).

The group also proposed to pare down the list of over 40 indicators to a more parsimonious list of the 10–15 key indicators for reasons of practicality at the field level. However, in subsequent meetings of the SDWG, it was judged premature to reduce the list of indicators without the benefit of field testing. Thus, the full list of indicators is presented here.

Since a number of indicators could be measured at two or even three levels, it was decided to identify the level at which each indicator was most important. In most cases this was the client level. If results were satisfactory at the client level, one could assume adequate performance at the manager/provider levels. If deficiencies were found, one would then move back, first to the provider level, then to the manager level (or enabling systems) to identify the source of the deficiency in an effort to remedy the problem.

In the following list, the indicators are categorized according to the six elements of the Bruce Framework on Quality of Care. Where appropriate, they are presented (within element) in chronological sequence (i.e., the order in which they would be expected to happen in a service delivery setting).

For consistency, the indicators have been worded in the positive sense (e.g., client receives his/her method of choice). However, the instruments to be developed from these indicators should use neutral wording that gives equal weight to a positive or negative response.

The right–hand column indicates the type(s) of data collection approach(es) that can be used to obtain data for each indicator. The codes are as follows:

AR = administrative (program) records

CR = client record review

CS = client survey

EI = exit interview with client

FG = focus group

OB = observation (client–provider interaction, clinical procedure, etc.)

PS = provider survey

INDICATORS OF QUALITY OF CARE IN CLINIC–BASED PROGRAMS

Element/Indicator	Data Collection Approach(es)
Interpersonal Relations	
1. Service providers are trained in interpersonal relations	AR, PS
2. Provider establishes rapport for assessing personal situation (family circumstances, nature of sexual relationships)	OB, CS, EI, FG
3. Client reports feeling:	CS, EI, FG
a. welcomed by staff	
b. at ease asking questions	
c. treated with respect/politeness by providers	
Choice of Method	
4. Number of methods approved for use at the SDP	AR
5. Number/range of methods available at SDP[1]	OB, AR
6. Provider offers all appropriate methods	OB, EI
7. Provider places no unnecessary restrictions on method choice	PS, AR
8. Client receives her/his method of choice[2]	CS, EI
9. Provider refers client to an existing, accessible site for methods unavailable at SDP	PS, OB, EI
Information Given to Clients	
10. Provider demonstrates good counseling skills (e.g., providing information, eliciting information, answering questions)	OB, EI
11. Provider has checklist available on information to cover during counseling session	OB, PS
12. Provider gives accurate and unbiased overview of all methods	OB, EI
13. Provider gives accurate, relevant[3] information on method accepted:	OB
a. how to use	
b. advantages and disadvantages	
c. side effects (balanced presentation)	
d. primary and secondary precautions	

[1] List of all methods physically available at the service delivery point (SDP) on the day of data collection. The indicator should be interpreted taking into account the number of methods approved for the country and appropriate to the type of SDP.

[2] This indicator must be interpreted in connection with indicator #34. Program should not be penalized for withholding the desired method if it is not medically appropriate.

[3] "Relevant" information will differ for new users, who need complete information, and continuing users, who may need little or no information.

Element/Indicator	Data Collection Approach(es)
e. complications that require referral	
f. resupply	
g. other important information[4]	
14. Provider asks client to repeat key information on method chosen (how to use, side effects, what to do if they occur, etc.)	OB, CS, EI
15. Client correctly explains method chosen:[5]	CS, EI
a. how to use	
b. possible side effects	
c. what to do if side effects occur	
d. when to return	
e. where to return	
16. Informational materials are available (printed, model, sample, etc.) on specific methods	OB
17. Privacy is acceptable for:	OB, EI
a. counseling	
b. exam (if any)	
18. Consent form is available and signed by client (VSC)	OB, CR
Technical Competence	
19. Written guidelines on FP practice are available at SDP	AR
20. Job descriptions exist for each position	AR
21. Formal mechanisms exist to review/screen potential service providers	AR
22. Education/training criteria exist for service tasks	AR
23. New staff are trained regarding institution's guidelines	AR, PS
24. Clinical providers have received training relevant to the job	AR, PS
25. All staff receive periodic refresher/in–service training	AR, PS

[4] To be determined at program level (e.g., the pill doesn't protect against HIV infection).

[5] Experience indicates that it may be difficult for interviewers to correctly record and assess the adequacy of responses given by clients. Although #13 and #15 are similar, a poor response on #15 does not necessarily mean that the provider's explanation was inadequate (#13); thus, both are retained as indicators.

Element/Indicator	Data Collection Approach(es)
26. Basic items are present for delivering methods available at SDP:[6]	OB, PS
a. sterilizing equipment	
b. gloves	
c. blood pressure	
d. specula	
e. adequate lighting	
f. water	
27. Provider can accurately explain contraception:[7]	PS
a. how to use	
b. advantages and disadvantages	
c. side effects (balanced presentation)	
d. primary and secondary precautions	
e. complications that require referral	
f. resupply	
28. Provider demonstrates skill at clinical procedures (according to guidelines)	OB
29. Provider demonstrates ability to recognize/identify contraindications (consistent with guidelines)	OB, PS
30. Provider avoids tests, examinations, and waiting periods that are not medically justified[8]	OB, AR, PS
31. Provider follows infection control procedures (outlined in guidelines)	OB
32. All levels of service providers receive routine supervision:	AR, PS
a. regular	
b. useful (e.g., providers' knowledge and clinical skills monitored)	
33. SDP is capable of handling HIV, other STDs, and reproductive tract infections (RTIs):	AR, PS
a. identification	
b. diagnosis	
c. referral	
d. prevention counseling	
e. treatment and counseling	

[6] Not all supplies are required for the delivery of all methods.

[7] "Can explain" refers to the ability to provide correct answers on a knowledge test. This is different from actually providing these explanations on the job (see indicator # 13).

[8] Based on local service delivery guidelines.

Element/Indicator	Data Collection Approach(es)
34. Client receives an appropriate method:	OB, CR
a. medically appropriate	
b. appropriate for sexual lifestyle (including risk of STDs and HIV)	
Mechanisms to Ensure Continuity [9]	
35. Provider encourages client to return as needed	OB, EI, CS, FG
36. Follow–up/return schedule is appropriate/reasonable	CS, CR
37. Client can obtain resupplies easily	OB, CS, EI
a. supply of all methods offered at SDP is adequate	
b. system for resupply is reliable (prevents stockouts)	
38. Clients past–due for follow–up are identified	AR, CR
39. Clients past–due for follow–up are contacted	AR, CR
40. Reasons for non–return are identified	CS, FG
Appropriateness and Acceptability of Services [10–11]	
41. Clients and non–users perceive that:	CS, EI, FG
a. privacy/confidentiality for counseling is acceptable	
b. privacy/confidentiality for exam is acceptable	
c. waiting time is acceptable	
d. time with provider is acceptable	
e. hours/days are convenient	
f. staff is acceptable in terms of gender, ethnic group, age	

[9] Indicators #13–c, 13–e, 13–f, 15–b, 15–d, 15–e are also relevant to this element.

[10] Some would argue that "physical access" to the facility influences how acceptable it is to the client. However, under The EVALUATION Project, we have treated "access to services" as a separate (i.e. independent) dimension of the supply environment and thus have not included it on this list.

Another possible indicator of quality is affordability. However, this indicator has been excluded from this list on the grounds that the key dimension is not affordability per se, but value, which takes into account both quality and cost.

[11] Note: a client questionnaire on attitudes toward the service should end with an open–ended question on other aspects of service that could be improved.

Element/Indicator	Data Collection Approach(es)
42. Clients and non–users perceive facility to be adequate in terms of:	CS, EI, FG
a. waiting room	
b. exam room	
c. cleanliness/hygiene	
d. water	
e. toilet facilities	
f. _____ (other)	
OUTCOMES	
Number of new acceptors/users	AR
Complication rate for specific methods	CR,CS
Correct, consistent use of temporary methods (Continuation rate – of any method)	CR, CS, FG, EI
Number of new clients recommended by other users	CR, CS, EI
Number of users that recommend service to someone else	CS, EI
Percent of clients that achieve reproductive intentions	CS, EI

Indicators for Evaluating Operations Research[1]

Contractor Compliance with the Terms of the Contract

- Number of studies
- Types of countries (e.g., priority versus non–priority)
- Number of reports, workshops, participants trained

Process of Developing the OR Agenda

- Consultation with OR "constituencies:"[2]
  - USAID Mission in country
  - USAID/W CTO
  - Others in Office of Population
  - USAID/W Bureaus
  - Regional AID offices (e.g., REDSO)
  - CAs in the U.S.
  - CAs in the field
  - Own institution
  - Host country government (MOH)
  - NGOs in country
  - Other donors (UNFPA, UNICEF, WHO, etc.)
  - Advocacy groups (e.g., women's groups)
  - Scientific community
- Collaborative nature of the process
- Perceived utility of the OR agenda by the constituency groups

Quality of the Research

Study Design

- Ability to answer the research question
  - Design in its original form
  - Design as actually implemented
- Balance between available resources and robust design
- Potential impact on policy
- Research question relevant to the OR agenda (in–country and globally)

---

[1] The above list represents the product of the Operations Research Working Group Meeting held in October 1993. It constitutes an update of the preliminary list of indicators developed during the first meeting of the group (The EVALUATION Project, 1992d).

[2] Choice of constituency groups depends in part on the topic of the OR study.

- Potential for replication
- Evidence of improvements in research methods and tools

Conduct of the Research

- Implementation and quality control of fieldwork
- Timely monitoring and oversight of field activities
- Quality of data analysis and presentation
- Evidence of effort to solve unanticipated events

Dissemination of Results

- Range of channels (seminars, workshops, presentations, newsletters, publications in peer–reviewed journals, etc.)
- Quantity of dissemination activities
- OR constituencies reached
- Timeliness

Utilization of Results

- Evidence that OR staff work with local program managers and staff of other CAs to utilize information for program improvement
- Changes implemented as a result of OR study:
  - in policy
  - in service delivery procedures:
    - at same site
    - in expanded program (upscaling)
  - in programs of other organizations:
    - in–country
    - elsewhere
- Information used as basis of further research
- Replication of methodology elsewhere
- Utilization of results or methodology in related fields

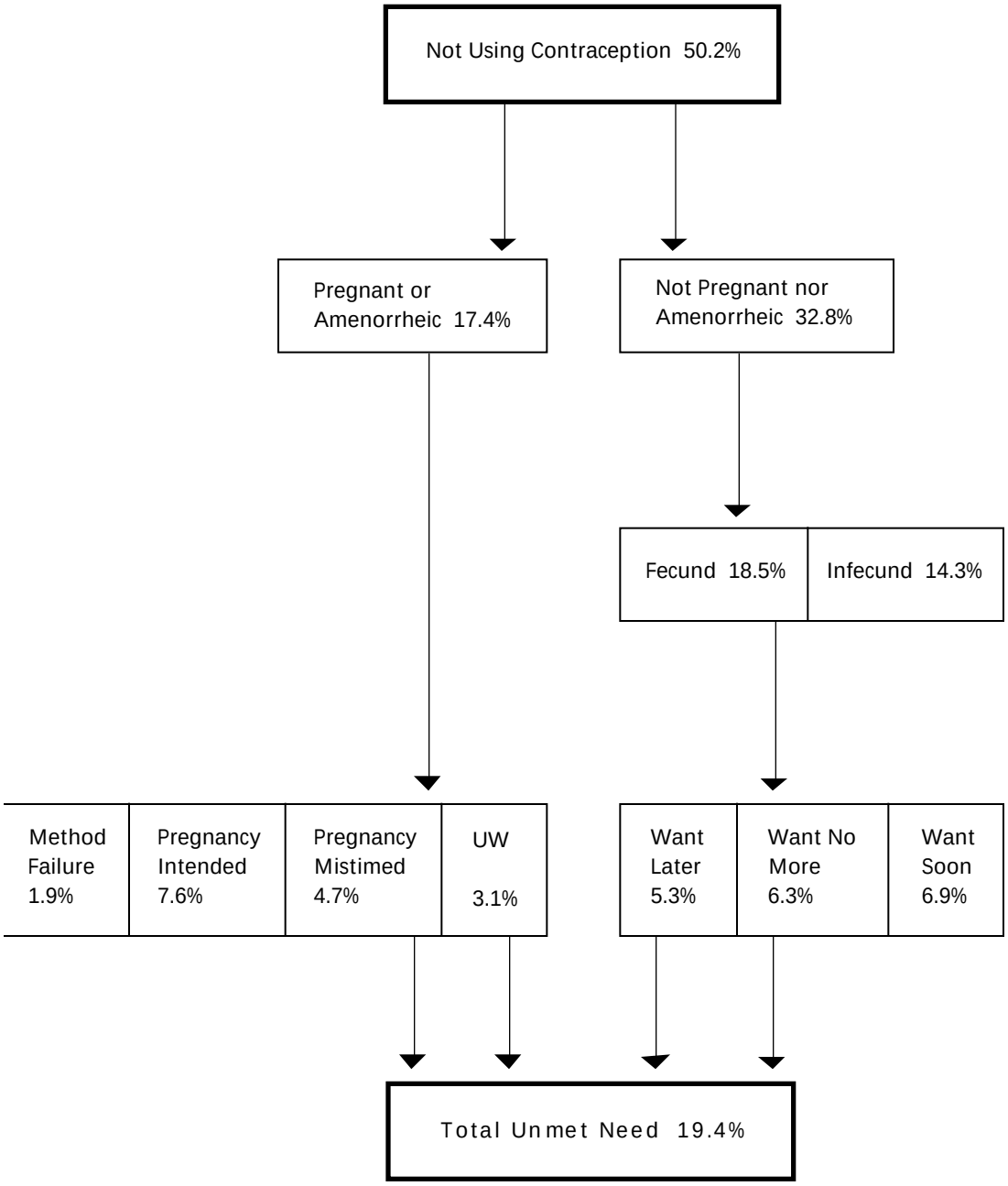
Institutionalization of OR as a Management Tool

- Increased receptivity of program managers to OR as a tool
- Orientation of managers toward a proactive approach
- Increased research capability among local researchers (in research design, data collection, processing, analysis, report preparation)
- Strengthening of research infrastructure (computers, reference documents, VCRs, software, link to INTERNET)

Other

- Diversity of skills that OR staff, consultants, and collaborators bring to the issue

Illustration of the Calculation of Unmet Need for Family Planning from the Dominican Republic DHS



Source: Westoff and Ochoa (1991)

Illustrative Tabulation Showing the Categories of Survey Respondents Required to Calculate Demand for Family Planning and Unmet Need – DHS Data, Selected Countries (see Chapter VI)

			Pregnant or Amenorrheic				Fecund				
Country	Infecund	Current Users	Method Failure	Intended Pregnancy	Mistimed Pregnancy	Unwanted Pregnancy	Want Child Soon	Want Child Later	Want No More Children	Total	Number of Women
Sub–Saharan Africa											
Botswana	15.8	33.0	1.7	12.9	8.8	0.9	9.8	10.6	6.5	100.0	1708
Burundi	13.1	8.7	NA	45.6	10.4	2.3	7.6	7.3	5.1	100.0	2669
Ghana	12.3	12.9	NA	27.4	10.7	2.0	12.2	15.6	7.0	100.0	3156
Kenya	13.3	26.9	NA	15.9	12.0	4.0	5.9	10.5	11.5	100.0	4765
Liberia	18.2	6.4	NA	28.4	1.7	8.4	14.1	18.2	4.6	100.0	3529
Mali	19.3	4.7	NA	38.0	4.2	1.4	15.2	13.0	4.3	100.0	2948
Togo	9.1	12.1	NA	27.2	11.7	3.1	11.4	16.7	8.5	100.0	2454
Uganda	15.5	4.9	NA	33.4	11.7	2.1	19.0	8.2	5.2	100.0	3180
Zimbabwe	8.2	43.1	NA	14.6	5.0	3.5	12.4	5.1	8.1	100.0	2643
North Africa											
Egypt	15.4	37.8	1.8	13.6	3.6	5.2	6.2	6.5	9.8	100.0	8221
Morocco	15.5	35.9	2.8	15.5	6.2	3.1	8.1	6.2	6.5	100.0	5447
Tunisia	10.5	49.8	1.6	13.3	4.6	3.3	5.1	6.0	5.8	100.0	4012
Asia											
Indonesia	19.9	47.8	0.9	11.5	2.7	0.9	3.9	7.4	5.0	100.0	0907
Sri Lanka	11.4	61.7	2.0	7.4	2.3	1.0	5.3	4.8	4.1	100.0	5442
Thailand	12.3	65.5	0.5	6.7	1.8	1.1	3.9	3.8	4.3	100.0	6236
Latin America/Caribbean											
Bolivia	17.6	30.3	3.8	9.4	6.3	11.0	3.2	3.2	15.2	100.0	4941
Brazil[1]	6.6	66.2	2.2	6.9	2.6	2.3	5.4	2.2	5.7	100.0	3471
Colombia	8.2	64.8	2.7	6.5	3.3	2.1	4.4	1.8	6.2	100.0	2850
Dominican Republic	14.3	49.8	1.9	7.6	4.7	3.1	6.9	5.3	6.3	100.0	4133
Ecuador	12.1	44.3	2.3	12.8	4.4	2.9	4.4	6.3	10.5	100.0	2957
El Salvador	13.5	47.3	0.6	8.2	3.7	4.0	4.5	10.2	8.1	100.0	3164
Guatemala[1]	10.4	23.2	0.9	30.1	5.8	4.0	6.1	10.6	9.0	100.0	3377
Mexico	11.0	52.7	2.2	8.6	4.0	3.8	1.5	7.0	9.3	100.0	5662
Peru	12.4	45.8	4.3	7.0	5.2	6.5	2.8	2.9	13.1	100.0	2900
Trinidad & Tobago	18.6	52.7	2.2	4.5	1.9	1.2	5.9	6.3	6.7	100.0	2617

Note: The ratios of intended to mistimed pregnancies for currently pregnant and amenorrheic women in Mexico and Zimbabwe were estimated from regional figures.

NA: Not applicable

[1] Figures are for women 15–44

Source of data:Westoff and Ochoa (1991)

The
EVALUATION
Project

PROJECT FUNDED BY USAID

Contract Number: DPE-3060-C-00-1054-00

THE EVALUATION PROJECT

Carolina Population Center
University of North Carolina at Chapel Hill
CB 8120 University Square
Chapel Hill, NC 27516-3997 USA

COLLABORATING INSTITUTIONS

The Futures Group International
1050 17th Street NW, Suite 1000
Washington, DC 20036 USA

Tulane University
Center for International Health and Development
School of Public Health and Tropical Medicine
1440 Canal Street, Suite 1300
New Orleans, LA 70112-2737 USA