

# **MONITORING PROGRESS TOWARD THE GOALS OF THE WORLD SUMMIT FOR CHILDREN**



**A PRACTICAL HANDBOOK FOR MULTIPLE-INDICATOR SURVEYS**

**Planning Office  
Evaluation and Research Office  
Programme Division**

---

**United Nations Children's Fund**

New York  January 1995

Copyright © 1995  
United Nations Children's Fund  
3 UN Plaza  
New York, NY 10017  
USA

January 1995

Permission to translate or reproduce this document, in part or in whole, will normally be granted freely by the UNICEF Representative when reproductions or translations are not intended for sale or for use in conjunction with commercial purposes.

The designations employed in this publication and the presentation of the material do not imply on the part of the United Nations Children's Fund the expression of any opinion whatsoever concerning the legal status of any country or territory, or of its authorities, or the delimitations of its frontiers.

*All correspondence should be addressed to:*

ACTION  
Planning Office, UNICEF  
3 UN Plaza  
New York, NY 10017  
Fax: (212) 303-7959

# CONTENTS

<i>Preface</i> .....	ix
<i>Index of Boxes, Tables and Figures</i> .....	xi
<b>CHAPTER 1: INTRODUCTION</b> .....	<b>1.1</b>
The Mid-Decade Goals: Background .....	1.1
Measurement of Progress .....	1.2
Reporting at Mid-Decade on the Mid-Decade Goals .....	1.4
The Mid-Decade Review .....	1.4
Monitoring Systems for National Reports on the Mid-Decade Goals .....	1.5
Development of Appropriate Monitoring Indicators .....	1.5
Choosing a Monitoring Tool .....	1.6
What Are the Tools for Providing Information on These Indicators? .....	1.7
Choosing to Monitor Progress with a Survey .....	1.9
Deciding to Conduct a Survey: Is There a Need? .....	1.9
Who Is This Handbook For? .....	1.12
National Capacities: Today and Toward the Year 2000 .....	1.14
<b>CHAPTER 2: DECISIONS TO MAKE BEFORE STARTING</b> .....	<b>2.1</b>
Identifying and Contacting Potential Resource Persons .....	2.1
Deciding on the Level of Aggregation of the Indicators .....	2.2
Estimating How Long the Survey Will Take .....	2.2
Calculating How Much the Survey Will Cost .....	2.3
<b>CHAPTER 3: DESIGNING THE QUESTIONNAIRE</b> .....	<b>3.1</b>
Which Goals Can Be Monitored with a Survey? .....	3.1
What Are the Indicators of These Goals and What Information is Required to Measure Them? .....	3.2
What are the Requirements of a Good Questionnaire? .....	3.10
Decide How to Ask the Questions .....	3.14
Adapting and Translating Questionnaires .....	3.14
Planning the Data Management and the Analysis .....	3.15
Translating the Question Modules .....	3.15
Pretesting the Questionnaire .....	3.16
Decide Who the Respondents Will Be .....	3.18
The Model Questionnaire .....	3.19
Organization of the Questionnaire .....	3.19
Model Questionnaire Modules .....	3.23
Household Module .....	3.25 (Q1)

Water and Sanitation Module .....	3.26 (Q2)
Salt Iodization Module .....	3.26 (Q2)
Education Module .....	3.27 (Q3)
Modules for Mothers .....	3.28 (Q4)
Tetanus Toxoid (TT) Module .....	3.28 (Q4)
Care of Acute Respiratory Illness Module .....	3.29 (Q5)
Modules for Children Under Five Years of Age .....	3.30 (Q6)
Diarrhoea Module .....	3.30 (Q6)
Vitamin A Module .....	3.31 (Q7)
Breastfeeding Module .....	3.32 (Q8)
Immunization Module .....	3.33 (Q9)
Alternative Immunization Module .....	3.34 (Q10)
Anthropometry Module .....	3.35 (Q11)
Mortality Module .....	3.36 (Q12)
<b>CHAPTER 4: CHOOSING THE SAMPLE .....</b>	<b>4.1</b>
Understanding the Basic Concepts of Sampling .....	4.1
Simple Random Sampling .....	4.3
Systematic Sampling .....	4.3
Cluster Sampling .....	4.3
Sampling with Probability Proportional to Size .....	4.4
Procedure for Sampling with Probability Proportional to Size .....	4.5
Stratified Sampling (Stratification) .....	4.7
Implicit Stratification .....	4.8
How to Calculate the Sample Size .....	4.8
Decide on the Total Number of Clusters .....	4.18
Use a Flow Chart to Choose Your Sampling Scheme .....	4.19
How to Select the Large Areas .....	4.20
How to Select the Small Areas .....	4.22
How to Select the Clusters .....	4.23
Enumeration. ....	4.24
Segmentation. ....	4.25
Random Walk Methods. ....	4.25
Country Examples .....	4.26
Bangladesh .....	4.26
Brazil .....	4.27
Jordan .....	4.28

---

<b>CHAPTER 5: PREPARING FOR DATA COLLECTION</b> .....	<b>5.1</b>
Making Logistics Arrangements .....	5.1
Setting Up Central and Local Headquarters .....	5.1
Contacting Local Authorities .....	5.1
Deciding on Size and Composition of Field Teams .....	5.2
Arranging Transportation, Accommodations and Security .....	5.3
Obtaining and Preparing Copies of Local Maps .....	5.4
Preparing the Questionnaire .....	5.4
Pretesting the Questionnaire .....	5.5
Selecting the Field Workers .....	5.6
Choosing and Preparing the Equipment .....	5.8
Weighing Scales .....	5.9
Length/Height Boards .....	5.10
Salt Iodization Kits .....	5.10
Training the Field Workers .....	5.10
Briefing the Drivers .....	5.13
Carrying Out the Pilot Study .....	5.13
Setting Up Computers and Hiring Data Processing Staff .....	5.14
Making Arrangements for Returning the Questionnaires to Headquarters .....	5.15
<b>CHAPTER 6: CONDUCTING THE FIELD WORK</b> .....	<b>6.1</b>
How to Map and Locate the Households .....	6.1
Using Segmentation to Identify the Clusters .....	6.1
The Random Walk Method .....	6.5
Selecting the Starting Point on the Boundary Map .....	6.5
Using the EPI Method for Selecting the Starting Point .....	6.8
Using the Random Walk Method for Selecting Households .....	6.8
Using Different Sampling Schemes for Urban and Rural Areas .....	6.10
How to Handle Special Situations in the Field .....	6.11
What Each Interviewer Should Do .....	6.12
What Each Supervisor Should Do .....	6.13
<b>CHAPTER 7: PROCESSING THE DATA</b> .....	<b>7.1</b>
Introduction .....	7.1
Entering the Data into the Computers and Producing "Clean" Data Files .....	7.2
Data Entry .....	7.2
Data Checking .....	7.4
Evaluating the Quality of the Survey Data .....	7.5
Calculating the Indicators .....	7.6

Estimating the Margin of Error for Each Indicator .....	7.8
Calculating Weighting Factors for Non-Self-Weighting Samples .....	7.9
Sex- and Age-Specific Estimates .....	7.11
<b>CHAPTER 8: REPORTING AND USING THE SURVEY FINDINGS .....</b>	<b>8.1</b>
Reporting at Mid-Decade .....	8.1
Presenting Findings to the General Public .....	8.1
Producing a Preliminary Report .....	8.2
The Full Survey Report: Presenting Findings to Decision-Makers at Different Levels .....	8.3
The Executive Summary .....	8.3
Producing the Technical Report .....	8.5
Report the Indicator Estimates .....	8.6
From Data to Information .....	8.8
Using the Bangladesh Survey Data .....	8.8
The Final Report: Monitoring the Goals .....	8.9
Other Reporting Formats: Advocacy and Policy Dialogue .....	8.9
Using the Survey Results for Expanded Action .....	8.10
Subnational Planning: Looking Beyond National Averages .....	8.10
Interagency Cooperation .....	8.10
UNICEF Country Programme of Cooperation: Increased Efficiency and Effectiveness .....	8.11
Staff Assignment and Budget Exercise .....	8.11
Reviews of Progress by Heads of State, Mayors and the Public .....	8.11
<b>APPENDIX 1: Instructions for Interviewers .....</b>	<b>A1.1</b>
Introduction .....	A1.4
How to Handle an Interview .....	A1.4
General Points .....	A1.5
How to Fill in the Questionnaire .....	A1.6
Household Module .....	A1.6
Water and Sanitation Module .....	A1.9
Salt Iodization Module .....	A1.9
Education Module .....	A1.9
Modules for Mothers .....	A1.10
Tetanus Toxoid (TT) Module .....	A1.10
Care of Acute Respiratory Infection Module (optional) .....	A1.11
Modules for Children under Five Years of Age .....	A1.11
Diarrhoea Module .....	A1.11
Vitamin A Module .....	A1.12
Breastfeeding Module (optional) .....	A1.13

---

Immunization Module .....	A1.14
Alternative Immunization Module .....	A1.14
Anthropometry Module .....	A1.14
Mortality Module .....	A1.16
Filling in the Cluster Control Sheet .....	A1.17
<b>APPENDIX 2: Anthropometric Techniques .....</b>	<b>A2.1</b>
<b>Annex I, Summary Procedures: How to Weigh and Measure Children .....</b>	<b>A2.5</b>
Procedures and Precautions Before Measuring .....	A2.7
Layout of the Procedures .....	A2.7
Two Trained People Required .....	A2.7
Measuring Board and Scale Placement .....	A2.7
Age Assessment .....	A2.7
When to Weigh and Measure .....	A2.7
Weigh and Measure One Child at a Time .....	A2.8
Control the Child .....	A2.8
Coping with Stress .....	A2.8
Recording Measurements and Being Careful .....	A2.9
Strive for Improvement .....	A2.9
Nutritional Status Measurement: Summary Procedures .....	A2.10
Child Height Summary Procedure .....	A2.10
Child Length Summary Procedure .....	A2.12
Child Weight Summary Procedure .....	A2.14
Child Mid-Upper Arm Circumference Summary Procedure .....	A2.16
<b>APPENDIX 3: Organization and Administration of Data Processing .....</b>	<b>A3.1</b>
Managing the Data Base .....	A3.4
Record-Keeping .....	A3.4
How to Install and Use the Data Entry and Analysis Programs .....	A3.6
EPI INFO Package for Mid-Decade Goal Surveys .....	A3.6
Installing the Programs .....	A3.6
Running the Program .....	A3.6
Data Entry .....	A3.7
Notes on the specific modules .....	A3.8
Consistency Checks .....	A3.8
Variable Frequencies .....	A3.8
Data Analysis .....	A3.8
Data Back-Up .....	A3.9
Enter EPI INFO .....	A3.9

Quit to DOS Sub-directory .....	A3.9
Running CSAMPLE to Obtain Standard Errors for Cluster Surveys .....	A3.9
Notes to the Programmer .....	A3.11
Removing Modules .....	A3.11
Adding Modules .....	A3.11
Changing the Wording of Questions or Response Categories .....	A3.11
Changing the Analysis/Consistency Checks .....	A3.12
Merging Data Sets from Several Computers .....	A3.12
<b>APPENDIX 4: Definitions of the Indicators .....</b>	<b>A4.1</b>



# CHAPTER I

## INTRODUCTION

This handbook is intended to assist UNICEF staff in their work with national counterparts to measure progress toward the Mid-Decade Goals. It is a working document and as such has sections that may be removed and reproduced for specific purposes.

The handbook is a practical, immediate response to requests for guidance as the end of 1995—mid-decade—comes closer. In 1990, at the World Summit for Children, 71 heads of state and government and 88 senior officials from countries around the world pledged themselves to a Declaration and Plan of Action for Children. Based now on the commitments of 158 heads of state, countries around the world have proceeded to develop their own National Programmes of Action (NPA) for Children. Today these National Programmes of Action encompass 9 out of 10 children on earth.

### **THE MID-DECADE GOALS: BACKGROUND**

The Mid-Decade Goals emerged from the first two years of experience, beginning in 1990, with National Programmes of Action. By 1992–93, some of the goals for the year 2000 appeared to be within the reach of many countries by the end of 1995, if the world community would make a concerted effort to that end. Making a measurable difference for children right away, long before the year 2000, has particular urgency. Children and their parents have good reason to be suspicious about promises for improvements in their lives that take 10 years to accomplish.

Accordingly, regional consultations built consensus in 1992 and 1993 around a set of mutually shared goals that could be achieved by mid-decade. At the heart of the process was the Consensus of Dakar, reached by all governments of Africa meeting under the aegis of the Organization for African Unity (OAU). In 1992 the International Conference on Assistance to African Children adopted a set of 13 Mid-Decade Goals, which subsequently received summit-level endorsement by the OAU. While other regional consultations agreed upon a more ambitious set of goals for mid-decade, and have set out to meet them, the real challenges for improving the lives of children and women are in Africa. The Mid-Decade Goals which emerged from Dakar have come to form the moral minimum that all countries need to achieve by the end of 1995 as stepping stones to the goals for the year 2000,<sup>1</sup> and it is those goals that are addressed in this handbook.

---

<sup>1</sup>For fuller background, see *Report of the Secretary-General: Implementation of the General Assembly Resolution 45/217 on the World Summit for Children*, A49/326, 46th Session of the General Assembly (New York: United Nations, 1994).

## MEASUREMENT OF PROGRESS

Governments that signed the Declaration and Plan of Action for Children also committed themselves to monitoring progress toward the goals and objectives set for the year 2000 in the Plan of Action of the World Summit for Children. The Plan of Action called for each country to "establish appropriate mechanisms for the regular and timely collection, analysis and publication of data required to monitor social indicators related to the well-being of children" (para 34 [v]). Many governments have taken serious steps to do this in their own National Programmes of Action. Measurement of these indicators is an essential part of the process—both for providing information for action and for recognizing improvements. But guidance for doing so has lagged behind demand.

The Plan of Action requests "the assistance of the United Nations . . . to institute appropriate mechanisms for monitoring the implementation of the Plan of Action, using the existing expertise of the relevant United Nations statistical offices, the specialized agencies, UNICEF and other United Nations organs." The annual publication *The Progress of Nations* is one of UNICEF's contributions to that effort. In addition to highlighting how the measurement of progress lags behind action, *The Progress of Nations* has raised the following question: Does the absence of data on the well-being of children reflect an absence of real commitment? Surely, if the vagaries of world financial markets can be reported on a daily basis, then data on children should not be so woefully out-of-date or altogether lacking.

The Mid-Decade Goals have narrowed the information needs to specific areas critical to the survival and development of each child: the right to be protected against the immunizable diseases, to be treated when sick with diarrhoea, to be breastfed, to be well nourished, to be protected against the disorders of iodine and vitamin A deficiencies, to go to school and to have access to water as well as sanitation facilities.<sup>2</sup> Each of the goals is framed within the available resources of each country and the international community.<sup>3</sup>

Yet, recent reviews of data reported by countries to international agencies show that measurement at the national level still lags behind progress and data voids persist. This is demonstrated in Table 1.1, which shows the number of years that have elapsed, on average, between the last national on-the-ground surveys and the year 1994.

---

<sup>2</sup>*Human Rights Fact Sheet No. 10*, Convention on the Rights of the Child, Article 24, ISSN 1014-5567 (Geneva: Centre for Human Rights, United Nations, 1990).

<sup>3</sup>Convention on the Rights of the Child, Article 4.

**Table 1.1 Average age of data on Mid-Decade Goals for developing countries**

Goal	Indicator	Age of data in years (by region)				
		Sub-Saharan Africa	South Asia	East Asia and Pacific	Middle East and North Africa	Latin America and the Caribbean
1	Immunization—DPT	1.3	1.4	1.2	1	1.1
	—Measles	1.3	1.4	1.2	1	1.2
	—Polio	1.3	1.4	1.2	1	1.1
	Tetanus toxoid for pregnant women	2.8	1.6	3.9	3	13.6
	TB immunization	1.3	1.4	1.2	4.4	1.9
5	Children receiving adequate vitamin A <sup>a</sup>	← ←	Only 4 of 99 countries have data.			⇒ ⇒
6	Iodized salt consumption <sup>a</sup>	← ←	Only 12 of 99 countries have data.			⇒ ⇒
7	Use of ORT <sup>b</sup> (pre-1993 definition)	2.9	1.3	2.6	2.6	1.3
	Use of ORT <sup>b</sup> : increased fluids and continued feeding <sup>a</sup>	← ←	Only 13 of 99 countries have data.			⇒ ⇒
11	Stunting	9.4	8.7	12.5	10.8	8.1
	Underweight	8.9	8.7	7.9	9.1	7.6
12	Children reaching grade 5	7	8.9	10.6	6.4	7.2
	Entering grade 1 at recommended age	10.6	15.0	12.9	7.5	7.3
	Gross primary school enrollment	5.3	3.9	4.9	3.9	3.3
	Net primary enrollment ratio	8.7	11	8.5	6.1	4.6
13	Access to safe drinking water	4.7	2.9	4.3	4.7	3.2
	Access to sanitary excreta disposal	6.2	5.0	4.3	6.4	4.3

**Note:** Where no data are available for the period since 1975, the elapsed time for a country indicator is set at 20 years.

<sup>a</sup>Only recently defined.

<sup>b</sup>Oral rehydration therapy.

## REPORTING AT MID-DECADE ON THE MID-DECADE GOALS

The Mid-Decade Goals offer an immediate opportunity to close the gap between national-level measurement and actual progress. Not only are the goals widely supported, but it is also generally agreed that progress can be measured in a timely fashion in data-rich and data-poor countries as part of programmatic action to reach the goals. UNICEF has summarized the guidance of the World Health Organization (WHO) for the health goals (which account for 11 of the 13 Mid-Decade Goals) and of UNESCO for the education goal in its *Technical Guidelines for Monitoring the Mid-Decade Goals*.<sup>4</sup>

The remaining mid-decade goal—universal ratification of the Convention on the Rights of the Child—is the most important one. It needs no measurement, only action on the part of national legislatures to ratify the Convention and see that the documents are deposited with the United Nations Legal Office. Of all the Mid-Decade Goals it is the most advanced: 164 countries around the world have now ratified the Convention. Universal ratification requires only 22 remaining countries. There is every reason to hope that by 1995 this Convention will become the first universally ratified treaty of the world community.

The Convention on the Rights of the Child is the foundation for all the goals set at the World Summit for Children, as stated in the Summit's Plan of Action: "The aspirations of the international community for the well-being of children are best reflected in the Convention on the Rights of the Child" (Introduction, para 4).

Each mid-decade goal is quantified and expresses a time-bound target for progress through 1995. The goals are summarised in Table 1.3 (see end of this chapter) and compared with their related goals for the year 2000. Measurable indicators—which all agree can be used feasibly and affordably—are linked to these targets as an integral part of programme delivery. Indicators for these goals serve as summary markers of progress; in some cases they are proxy measures for information that is lacking but is very difficult or costly to collect on a national basis.

## THE MID-DECADE REVIEW

An assessment of status on the Mid-Decade Goals is near at hand. The Plan of Action of the World Summit for Children calls upon "the Secretary-General of the United Nations . . . to arrange for a mid-decade review at all levels, of the progress being made toward implementing the Declaration and Plan of Action" (para 35 [iv]). This review necessarily encompasses all national and international goals for children, but it will put special emphasis on the smaller set of Mid-Decade

---

<sup>4</sup>UNICEF CF/PROG/IC/94-003, Programme Instructions, 15 March 1994.

Goals that emerged from Dakar, and it will draw on national and regional reviews of progress. Such reviews, led in many cases by the head of state, are already under way in different regions. The Secretary-General will make his own report to the United Nations on a special occasion in 1996 after all national governments have made their own reports.

### **Monitoring Systems for National Reports on the Mid-Decade Goals**

There are many different ways to go about monitoring indicators in each goal area. Different systems are summarized in UNICEF's *Technical Guidelines for Monitoring Progress Toward the Mid-Decade Goals*. Each is part of its own national context and has different uses. National Programmes of Action give *prima facie* evidence of the variety of monitoring mechanisms and capacities. Most National Programmes have set up interministerial coordinating bodies for ensuring that goals are monitored across ministries as well as at different levels of government so that current status may be reported in a timely fashion to policy makers.

### **Development of Appropriate Monitoring Indicators**

UNICEF, working with WHO and UNESCO, has developed *indicators* to measure progress toward these Mid-Decade Goals that permit cross-country comparisons. An indicator is the basic tool for measuring progress, using a commonly agreed-upon definition of a specific situation. Indicators define the data to be collected, so they should be relatively easy to measure and interpret, and should provide valid and reliable information about the objectives they are meant to measure. Ideally, an indicator should also provide useful information to improve programme operations.

The primary indicators have been selected because they can be collected as an integral part of activating national strategies to reach the Mid-Decade Goals. They form a subset of a larger number of indicators already

**Indicators are indirect or proxy measures of a situation used to monitor and evaluate progress.**

in use for monitoring programs. This subset of indicators to be measured at the national level can provide data that are meant to incite people and policy makers to action. But these data may also be used to monitor programmes at several different levels. The same information may be collected for use by national policy makers and for managing programmes within the health sector, but depending on its planned use, such information may be analysed, interpreted and presented in various ways.

*Technical Guidelines for Monitoring Progress Toward the Mid-Decade Goals* provides further background for these indicators, definitions and common data sources and references. These indicators measure the impact of programmes and generally require techniques, samples and measurement that can be reported only periodically—perhaps at intervals of several years.

The World Health Organization regularly reviews national reports on health, with its *Monitoring of Progress Toward the Goals of Health for All by the Year 2000*. Its latest monitoring of national reports collected data from governments in 1993. A review of those data concluded that many essential indicators are out-of-date or altogether lacking for reporting on the Mid-Decade Goals, further reinforcing initial conclusions drawn in *The Progress of Nations*, in both its 1993 and 1994 releases.

With only one year left until the end of mid-decade, existing data gaps must be filled, especially for the Mid-Decade Goals. WHO and UNICEF share the view that there are ways to do this that will reinforce existing national capacity and strengthen sustained improvements in measurement over the longer term.

In September 1994, the WHO-UNICEF Intersecretariat for the Joint Committee on Health Policy recommended that

Where data are lacking or out-of-date . . . WHO, UNICEF and other international agencies . . . collaborate across all sectors in assisting countries to monitor progress, using[,] wherever appropriate and feasible[,] multiple indicator surveys.<sup>5</sup>

UNICEF can assist countries in reporting on progress by helping to build in-country monitoring capacity and by identifying and investing in tools that will yield timely information. Priority should be given to methodologies that yield nationally representative results and can stand up to the careful scrutiny required for reporting. National statistical offices are key partners, as are the intersectoral agencies established in National Programmes of Action for monitoring progress toward the Summit goals.

As the information on age of data (Table 1.1) suggests, national capacity to monitor these indicators is still lacking in many areas. For monitoring progress at the national level (for use in advocacy, for overall policy making and for technical and managerial use), the measurement of indicators needs to be consistent, using agreed-upon definitions so that the measures can be used to compare countries and to monitor changes over time. Surveys can, in most settings, complement and strengthen existing reporting systems and fill information gaps for seven of the Mid-Decade Goals, which are discussed in chapter 3.

### **Choosing a Monitoring Tool**

The measurement tools to obtain data on these indicators must meet certain qualifications: the data they produce must be valid and reliable, sensitive to changes and specific to each indicator. The

---

<sup>5</sup>UNICEF Executive Directive, "Multiple Indicator Cluster Surveys for the Mid-Decade Goals," CF/EXD/1994-011 (14 November 1994).

indicators must be readily quantifiable, using agreed-upon definitions and reference standards. The data used to calculate each indicator should also be comprehensive, supplying representative national information and appropriate subnational breakdowns as required.

### ***What Are the Tools for Providing Information on These Indicators?***

*Population census data* are important sources of information on the total population, its age structure and its geographical distribution. This information is particularly useful in countries that most need data for setting priorities and for allocating limited resources, because they are also the countries most likely to have incomplete or insufficient information from the *registration of vital events* (births, deaths, marriages, etc.).

Census data can provide *denominators* for indicators; for example, the number of under-fives or the number of children of primary-school age can be estimated from census reports. However, censuses are not carried out by the health sector and usually take place only every 10 years. It can take up to three or more years to produce from census data the information needed by the health sector, so their usefulness for monitoring is limited.

*Routine health service data* are relatively cheap and easy to collect, but often those responsible for their collection at the local level are not taught how to use the data themselves, are often overburdened by demands to fill in forms and do not receive feedback on how these data are used. For these reasons, the quality of the data collected is often quite low.

Moreover, health service statistics are usually collected for administrative purposes, not for monitoring purposes. Information required on the proper denominators for the indicators are usually not available. Health service statistics can provide information on the number of patients seen, or the number of visits, but they rarely provide information on the entire population at risk, or the total population that is covered by the service. The "population at risk" (for example, the number of live births in a year) can be estimated from census data, if they are up-to-date, and then combined with health service information (for example, the number of children under one year of age immunized in a given year) to measure indicators if the appropriate data have been collected.

Helping to strengthen routine reports from the health information system, streamlining the amount of information to be collected and assuring that the population census data are tabulated and adjusted to provide appropriate information on denominators for selected indicators make up an important long-term strategy for building in-country monitoring capacity. These two sources—the health information system and the census—should be your first stop when looking for data to measure these indicators. However, the health services do not routinely obtain some of the information required for monitoring the Mid-Decade Goals.

To supplement health service statistics, epidemiological *surveillance sites* are an important data source. As long as the data are collected nationally, or in representative sites across the country,

they can provide useful information on indicators such as immunization, as well as on interventions delivered to certain regions or groups in a population, and they can be particularly useful in providing data on specific diseases. For example, a survey may be necessary to estimate the number of diarrhoea cases treated with oral rehydration salts (ORS), but a sentinel disease surveillance system may provide good data on seasonal patterns of diarrhoeal disease incidence, which can be used to adjust survey estimates of diarrhoea incidence. These systems may also collect information on the occurrence of specific diseases and conditions, such as vitamin A deficiency, goiter and acute respiratory infections, that are too costly to investigate in large-scale surveys.

*Household sample surveys* are the most widely used method of providing data on health and social indicators when these other sources of data are deficient. Even well-developed health information systems and good registers of vital events serve an important *complementary* role, supplementing routine service statistics with timely data, and with data not usually collected in censuses and at service delivery points (SDPs). Every industrialised country has special ongoing survey programmes, usually very extensive health interview surveys and surveys of disease morbidity and disability, to provide this supplementary information. Surveys are the best source of data on programme coverage, and on differentials in health indicators. They can provide breakdowns of information by regional, social or ethnic groupings which are very difficult to obtain from routine data sources.

No other source exists for some of the information required to monitor the Mid-Decade Goals. For example, the only way to obtain nationally representative information on diarrhoea cases treated with ORT, or on the number of households using iodized salt, or about the nutritional status of young children, is to do a survey, or to ensure that the required questions, measurements or tests are carried out in an ongoing survey programme. Data already collected in household surveys—such as the Demographic and Health Survey (DHS) programme—may be able to provide the baseline information for some indicators, provided the right questions were asked. If such a survey is currently being planned it may be possible to add a module of questions to measure specific indicators of the Mid-Decade Goals.

While large-scale household surveys are not usually carried out routinely, they do provide data that can complement routine statistics, gathered at key points in time. In countries where the health information system leaves many gaps, allocating some programme resources to routine monitoring surveys at the national level may be appropriate. These surveys can be relatively small in scale, because for those countries where data needs are greatest, it is often more important to know the *order of magnitude* for an indicator than to measure it very precisely. This means that a very large survey is often not necessary for providing important basic data at the national level.



### ***Choosing to Monitor Progress with a Survey***

Information gaps persist in almost all countries, as already observed. Without new surveys many of these gaps will remain through 1995, limiting any objective assessment of progress. As the end of 1995 approaches, limited options are available beyond simple surveys for closing existing data gaps. Good national estimates of goal indicators can be obtained using modifications of the cluster survey methodology developed by EPI and CDD programme reviews.<sup>6</sup> Properly conducted, these surveys can produce data that will meet the rigorous requirements for reporting on the Mid-Decade Goals. Linked to improved reporting in routine systems, simple rapid surveys can be used for measurement, management and advocacy. They can be used in data-poor countries, and can fill information gaps and improve programme delivery in data-rich countries.

Most countries are likely to find the multiple-indicator surveys<sup>7</sup> immediately useful both for goal monitoring and for influencing—at affordable cost—policy, programme design and resource allocations to social priority sectors. Moreover, most countries can take advantage of substantial national technical capacity developed as a result of extensive experience with cluster surveys. These household cluster surveys are population-based data-gathering systems. They provide a snapshot of the target population at a specific point in time and, if well planned, should give a profile of those most at risk—that is, households that fall outside routine service delivery and reporting systems. The snapshot they present is taken from a sample of households and estimates, and are thus subject to error. Good reports from these surveys make clear the degree of confidence to be placed in the results. Chapter 4 of this handbook tells how to design a survey sampling strategy appropriate to your country's needs and resources.

### **DECIDING TO CONDUCT A SURVEY: IS THERE A NEED?**

Every survey, no matter how simple, incurs costs. Requirements of sample size and proper conduct of the field work are demanding. UNICEF offices should review all the household surveys planned or in progress to see what new information on the Mid-Decade Goals they may produce within two years. Where a survey is in the planning stage, UNICEF can help to focus the survey on information gaps by participating in the survey design and by supporting data analysis that gives priority to reports on indicators of the Mid-Decade Goals. The indicators that can be measured by a national survey are discussed in chapter 3 of this handbook.

---

<sup>6</sup>EPI stands for the "Expanded Programme on Immunization." CDD stands for "Control of Diarrhoeal Diseases."

<sup>7</sup>When using an acronym to refer to multiple-indicator surveys, we use MICS (multiple-indicator cluster surveys) to avoid confusion with "management information systems" (MIS).

Before making a decision to do a survey to monitor the Mid-Decade Goals, be sure you have checked to see what other surveys have been done recently, or are planned for the next year. If they employ nationally representative samples, it may be possible to measure the indicators by asking that special tabulations be made from these recent surveys. If a survey is in the planning stage, you may be able to include modules of questions to measure the indicators for which routine data are lacking, and to obtain the cooperation of the survey team to produce the necessary tabulations quickly, as soon as data collection ends.

 **EXAMPLE:**

In Kenya, a national nutrition survey had been conducted in the first part of 1994. There was no need to include measures of the nutrition goals in the monitoring survey planned for the summer of 1994.

Once these other sources of data have been thoroughly mined, you will know what further information gaps exist. Where existing sources of data do not provide representative information needed to calculate an indicator, you may be able to collect it in a special survey. In countries with little data and no other surveys planned, there is good reason to work with governments to design a survey to provide baseline data, point to problems and assist in future assessments. Guidelines for modules to measure indicators of progress for EPI programmes and CDD programmes have proven to be cost-effective and useful tools.<sup>8</sup> These have formed part of the standard training for health programme managers for the past decade, and can provide rapid turnaround from fielding to reporting results.

UNICEF may undertake to support new surveys when:

- All recent data on the subject have been analysed and used.
- No one else is collecting further data or planning to do so in the near future.
- The method proposed is cost-effective and sustainable.
- National monitoring capabilities will be strengthened as a result.

Simple surveys should not be pressed to provide reliable subnational reports on these indicators, nor to report on change over time unless these requirements are clearly part of the survey design specifications at the outset.

---

<sup>8</sup> WHO, *The EPI Coverage Survey. Training for Mid-level Managers*, WHO/EPI/MLM/91.10, (Geneva: World Health Organization, 1991); and WHO, *Household Survey Manual. Diarrhoea Case Management, Morbidity and Mortality*, WHO/CDD/SER/86.2 Rev1(1989) (Geneva: World Health Organization, Diarrhoeal Disease Control Programme, 1989).

When a survey is well conducted, the results should stand up to close scrutiny by governments, other international organizations and communities. The data supplied by a carefully planned and well-conducted rapid survey should supply valuable information for helping communities and governments to understand and monitor their progress, and to plan for the future rationally.

---

---

**The quality of the data obtained in a survey depends on the proper design of the questionnaire, on the sampling strategy and on good training and supervision of suitable interviewers.**

---

---

To make sure that happens, careful planning for the survey is essential. This handbook brings together the best guidance the international community has to offer in doing that. The step-by-step guidance contained in the handbook for planning and conducting a multiple-indicator monitoring survey is intended to complement and reinforce existing systems—not to replace them.

The handbook contains a model of a well-designed survey instrument to use for obtaining the needed information. This standardised questionnaire is a product of a collaborative effort involving many participants, both within and outside UNICEF. It uses a series of question modules designed to provide data for most of the primary indicators of the Mid-Decade Goals. These question modules are found in chapter 3 of this handbook.

The Model Questionnaire can be easily adapted to specific country situations. If, for example, very good and current data already exist for an indicator in a particular country, the relevant module can be dropped. This handbook also offers the user an opportunity to select modules for briefer surveys to fill data gaps or to add to existing survey instruments where measurement plans are already well advanced. They may be combined with other questions that help to illuminate the determinants of current status. In addition to those designed to measure indicators of the Mid-Decade Goals, several other modules are provided, to be included or not at the discretion of the particular country offices and their counterparts. The survey can be implemented at reasonable cost in a variety of country situations.

Before the final decision to conduct a survey is made, you should be able to answer three important questions:

- *Why are you doing the survey?*
- *How do you expect to use the results?*
- *To whom, and at what level, will the report of results be addressed?*

Answers to these questions should help to ensure that the survey will provide useful information for monitoring goals, for influencing policy and programme design and for encouraging policy makers and programme managers to allocate resources to social priority sectors. What data are needed and how they may be used by policy makers, programme managers, communities and the

general public should inform all planning decisions.

Chapter 2 of this handbook contains a list of key information and decisions that need to be made before a survey can be undertaken. UNICEF country directors, programme managers, and national counterparts should read this section carefully before starting to plan a survey. It provides advice on what you can expect to gain from a survey and how long it will take to carry out the entire process, from planning through implementation to reporting your results. Chapter 2 also contains a checklist of items that will need to be budgeted for when doing a national survey.

---

---

**Make certain that all are clear about why you are doing the survey and how you will use the results. Then plan the presentation of findings along with planning of the survey itself.**

---

---

## WHO IS THIS HANDBOOK FOR?

For any country in which data are altogether lacking, multiple-indicator surveys using cluster sampling provide a means to measure status on most of the Mid-Decade Goals by mounting a rapid, nationally representative survey. The guidance in this handbook is set out to ensure that surveys may be designed and fielded in a way to yield results that will stand up to international scrutiny. The modules have limited aims, but the survey platform does not.

In all regions there is a reservoir of national managerial experience to mount the surveys that will contribute to these reports. In EPI alone, over 4,500 coverage surveys and reviews helped build national survey experience in the 1980s, especially in those countries with weaker monitoring systems. In Africa, 1,686 were completed between 1978 and 1991. No region lacks this experience. Most surveys have been managed by nationals. When supported with international assistance, nationals have been an integral part of all survey operations.

From the very start, wide consultation with all those who hold a stake in the results will enhance the survey's impact and acceptance of the findings. This means involving the Ministry of Health, perhaps the Planning Ministry and the Statistical Bureau, as well as nongovernmental organizations (NGOs) working for women and children. Widespread involvement, especially by key national training institutions, such as university public health programs or medical schools and university statistics departments, will also build capacity for future surveys and increase the long-term sustainability of monitoring efforts. All these different constituencies can be brought together to work toward achieving the commitments made by leaders at the World Summit for Children when the survey results are ready to report.

Different sections of the handbook should be read by different members of the survey team. Some sections should be translated, photocopied and given to field staff during their training. Table 1.2 can be used to guide country programme directors, national counterparts and members of the

survey team to the sections they should read.

**Table 1.2 Key to contents of the handbook**

<b>Section</b>	<b>Contents</b>	<b>Who should read</b>	<b>Need to translate?</b>
Chapter 1: Introduction	Monitoring progress at mid-decade Choosing a monitoring tool Deciding to conduct a survey	Programme director, national counterpart, survey coordinator	
Chapter 2: Decisions to Make Before Starting	Who will the technical resource persons be? What kind of data are required? How big should the survey be? How long will it take? What will it cost?	Programme director, national counterpart, technical resource persons, survey coordinator	
Chapter 3: Designing the Questionnaire	Which indicators will be assessed? What information is required? Target respondents for the different indicators Questionnaire modules	Survey coordinator	
Chapter 4: Choosing the Sample	Principles of sampling Critical decisions for sample size Sample size estimation Alternative sampling strategies	Survey coordinator, technical resource persons	
Chapter 5: Preparing for Data Collection	Logistical arrangements Preparing maps Printing questionnaires Selecting the interviewers Training field staff Arranging for data processing Pilot survey	Survey coordinator, technical resource persons	
Chapter 6: Conducting the Field Work	Choosing households to interview Carrying out interviews Supervising field operations	Survey coordinator, technical resource persons, field supervisors, interviewers	Yes (instruc- tional material)
Chapter 7: Processing the Data	Data entry Data checking Evaluating data quality Calculating the estimates	Survey coordinator, technical resource persons	
Chapter 8: Reporting and Using the Survey Findings	Presenting the findings Using the results	Programme director, national counterparts, survey coordinator, technical resource persons	

*(table continues on next page)*

*(Table 1.2, continued)*

Appendix 1: Instructions for Interviewers	How to fill in the questionnaire	Survey coordinator, technical resource persons, supervisors and interviewers	Yes
Appendix 2: Anthropometric Techniques	How to weigh and measure children	Survey coordinator, technical resource persons, interviewers, supervisors, measurers	Yes
Appendix 3: Organization and Administration of Data Processing	Managing the data base Running and installing the data entry and analysis programs Obtaining standard errors for cluster surveys Programmer notes	Technical resource persons	
Appendix 4: Definitions of the Indicators	Indicators measured with survey data	Technical resource persons	

### **NATIONAL CAPACITIES: TODAY AND TOWARD THE YEAR 2000**

Assistance from the United Nations Statistical Office, with support from UNFPA, and in consultation with other United Nations agencies, especially the World Health Organization and UNESCO, have helped to build on country experience to prepare this survey instrument for the Mid-Decade Goals. Under the guidance of the Planning and Coordination Office, all relevant technical clusters within UNICEF and the Evaluation and Research Office have also contributed to the question modules.

This handbook was developed in response to a clear demand by countries and field offices. In August of 1994, a global workshop on the multiple-indicator surveys was hosted by UNICEF Bangladesh in Dhaka, where an interagency team, supported by Bangladesh's Bureau of Statistics as well as by regional and international centers of excellence, prepared all UNICEF regional advisers in monitoring and evaluation to put in place regional support for such surveys.

When many gaps still exist in the information needed for monitoring, helping countries in the process of planning and carrying out a survey is an important way to strengthen national monitoring efforts for the future. That is why it is important to involve personnel from national institutions such as medical and public health schools, health worker training institutes, university statistics departments and social science departments. The effort put into building survey capacity

will increase over time when a pool of trained individuals is being developed.

Each country programme is well placed to assess where the strongest entry point may be to help national counterparts. The challenge is to bring together different sectors, as has been done in Bangladesh and Kenya, to collaborate in the cost-effective use of a shared tool.

This small window of opportunity will shortly close for those who postpone strategic work-planning for 1995. This handbook contains the technical guidance needed to assist in each step of this process, from deciding whether to do a survey to actually conducting and reporting on one.

#### **Box 1.1**

##### **WHAT MULTIPLE-INDICATOR SURVEYS CAN DO**

- Fortify local-level programme monitoring
- Satisfy national-level goal-monitoring needs
- Perform at low cost
- Produce rapid findings
- Strengthen existing national capacities for monitoring
- Ensure internationally comparable results

**TABLE 1.3**  
**WORLD SUMMIT FOR CHILDREN:**  
**A COMPARISON OF GOALS AND INDICATORS**  
**FOR THE YEARS 1995 AND 2000**

Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Indicators
<p>Maintenance of a high level of immunization coverage (at least 90% of children under one year of age by the year 2000) against diphtheria, pertussis, tetanus, measles, poliomyelitis, tuberculosis, and against tetanus for women of child-bearing age.</p>	<p>Proportion of children immunized against diphtheria, pertussis and tetanus (DPT3) before their first birthday.</p> <p>Proportion of children immunized against measles before their first birthday.</p> <p>Proportion of children immunized against poliomyelitis (OPV3) before their first birthday.</p> <p>Proportion of children immunized against tuberculosis before their first birthday.</p> <p>Proportion of pregnant women immunized against tetanus.</p> <p>Proportion of children protected against neonatal tetanus through immunization of their mother.</p>	<p>(1) Elevation of immunization coverage of six antigens of the Expanded Programme on Immunization to 80% or more in all countries.</p>	<p>All indicators as in column 2 (Year 2000 Goal Indicators).</p>
<p>Elimination of neonatal tetanus by 1995.</p>	<p>Annual number of cases of neonatal tetanus.</p>	<p>(2) Same as Year 2000 Goal.</p>	<p>Same as Year 2000 Goal Indicator.</p> <p>Proportion of districts reporting neonatal tetanus cases.</p>



Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Indicators
Reduction by 95% in measles deaths and reduction by 90% of measles cases compared with pre-immunization levels by 1995, as a major step to the global eradication of measles in the longer run.	Annual number of under-five deaths due to measles.  Annual number of cases of measles.	(3) Same.	Same.  Same.
Global eradication of poliomyelitis by the year 2000.	Annual number of cases of polio.	(4) Elimination of polio in selected countries and regions.	Same.  Proportion of districts reporting polio cases.
Virtual elimination of vitamin A deficiency, and its consequences, including blindness.	Proportion of children 2 to 6 years of age who are nightblind.  Proportion of children 6 months to 6 years of age with serum vitamin A below 20 micrograms/dl.  Proportion of lactating women with breast milk vitamin A below 30 micrograms/dl (or less than 8 micrograms/gram of milk fat).	(5) Virtual elimination of vitamin A deficiency. <i>(At least 80% of all children under 24 months of age in areas with vitamin A deficiency receive adequate vitamin A.)</i>	Proportion of children under 24 months of age receiving adequate vitamin A (in vitamin A-deficient areas).

Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Indicators
Virtual elimination of iodine deficiency disorders.	<p>Proportion of population in iodine-deficient areas consuming adequately iodized salt.</p> <p>Proportion of children 6 to 11 years of age with goitre of any grade (palpable and visible combined).</p> <p>Median value of the concentration of urinary iodine in school children should be greater than 10 micrograms/dl of urine.</p> <p>Proportion of newborns (cord blood or aged 3 days to 3 weeks) with serum TSH levels above 5 mIU/L.</p>	(6) Universal iodization of salt in IDD-affected countries.	Proportion of households consuming adequately iodized salt according to agreed criteria: (1) in the whole country and (2) in areas known to be at high risk of IDD.
Reduction by 50% in the deaths due to diarrhoea in children under the age of five years and 25% reduction in the diarrhoea incidence rate.	<p>Proportion of diarrhoea episodes in under-fives treated with ORT (increased fluids) and continued feeding.</p> <p>Annual number of under-five deaths due to diarrhoea.</p> <p>Average annual number of episodes of diarrhoea per child under five years of age.</p>	(7) Achievement of 80% usage of ORT (increased fluids) and continued feeding as part of the programme to control diarrhoeal diseases.	<p>Proportion of diarrhoea episodes in under-fives treated with oral rehydration salts (ORS) and/or recommended home fluids (pre-1993 ORT definition).</p> <p>Proportion of diarrhoea episodes in under-fives treated with ORT (increased fluids) and continued feeding.</p> <p>Proportion of population that has a regular supply of ORS available in their community.</p>

Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Indicators
Empowerment of all women to breastfeed their children exclusively for four to six months and to continue breast-feeding, with complementary food, well into the second year.	<p>Proportion of infants less than four months (120 days) of age who are exclusively breastfed.</p> <p>Proportion of children 20 to 23 months of age who are breastfeeding.</p> <p>Proportion of infants 6 to 9 months of age (180 to 299 days) who are receiving breast milk and complementary food.</p> <p>Proportion of all hospitals and maternity facilities which are "baby-friendly" according to global BFHI criteria.</p>	(8) Ending and preventing free and low-cost supplies of breast milk substitutes in all hospitals and maternity facilities. Having target hospitals and maternity facilities achieve "baby-friendly" status in accordance with BFHI global criteria.	<p>Proportion of hospitals and maternity facilities targeted for BFHI by end of 1995.</p> <p>Proportion of hospitals and maternity facilities that have been officially designated as "baby-friendly" in accordance with global criteria.</p>
Elimination of guinea-worm disease (dracunculiasis) by the year 2000.	<p>Annual number of cases of dracunculiasis in the total population.</p> <p>Number of villages with any cases of dracunculiasis.</p>	(9) Interrupt guinea-worm disease (dracunculiasis) transmission in all affected villages by the end of 1995.	<p>Same.</p> <p>Number of villages with any cases of dracunculiasis in the last year.</p>
		(10) Ratification of the Convention on the Rights of the Child (CRC) by all countries.	Depositing the Instruments of Ratification with the United Nations Legal Office.

Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Indicators
<p>Between 1990 and the year 2000, reduction of severe and moderate malnutrition among under-five children by half.</p>	<p>Proportion of under-fives who fall below minus 2 standard deviations from median weight for age of NCHS/WHO reference population.</p> <p>Proportion of under-fives who fall below minus 3 standard deviations from median weight for age of NCHS/WHO reference population.</p> <p>Proportion of under-fives who fall below minus 2 standard deviations from median height for age of NCHS/WHO reference population.</p> <p>Proportion of under-fives who fall below minus 3 standard deviations from median height for age of NCHS/WHO reference population.</p> <p>Proportion of under-fives who fall below minus 2 standard deviations from median weight for height of NCHS/WHO reference population.</p> <p>Proportion of under-fives who fall below minus 3 standard deviations from median weight for height of NCHS/WHO reference population.</p>	<p>(11) Reduction of 1990 levels of severe and moderate malnutrition by one-fifth (1/5) or more.</p>	<p>Same.</p> <p>Same.</p> <p>Same.</p> <p>Same.</p>

Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Indicators
<p>Universal access to basic education, and achievement of primary education by at least 80% of primary school-age children through formal schooling or non-formal education of comparable learning standard, with emphasis on reducing the current disparities between boys and girls.</p>	<p>Proportion of children entering first grade of primary school who eventually reach grade 5.</p> <p>Number of children enrolled in primary school who belong in the relevant age group, expressed as a percentage of the total number in that age group (net enrollment).</p> <p>Proportion of children of primary-school-entry age who enter grade 1 at that age.</p> <p>Proportion of children aged 10 to 12 years reaching a specific level of learning achievement in literacy, numeracy and life skills.</p>	<p>(12) Strengthen Basic Education so as to achieve reduction by one-third of the gap between: (a) primary school enrollment and retention rates in 1990 and universal enrollment and retention in primary education of at least 80% of school-age children, and (b) primary school enrollment and retention rates of boys and girls in 1990.</p>	<p>Same.</p> <p>Same.</p> <p>Same.</p> <p>Number of children enrolled in primary school expressed as a percentage of the total number of children of primary-school age (gross enrolment).</p> <p><i>(All above indicators by gender.)</i></p>
<p>Universal access to safe drinking water.</p>	<p>Proportion of population with access to an adequate amount of safe drinking water located within a convenient distance from the user's dwelling.</p>	<p>(13) Increase water supply and sanitation so as to narrow the gap between the 1990 levels and universal access by the year 2000 of water supply by one-fourth and of sanitation by one-tenth. <i>(It is important that the definitions pertaining to access, adequate amount, safe and sanitary facilities be reported on for the indicators.)</i></p>	<p>Same <i>(but expressed in absolute numbers and as a percentage of the total population).</i></p>
<p>Universal access to sanitary means of excreta disposal</p>	<p>Proportion of population with access to a sanitary facility for human excreta disposal in the dwelling or located within a convenient distance from the user's dwelling.</p>	<p>(13) Increase water supply and sanitation so as to narrow the gap between the 1990 levels and universal access by the year 2000 of water supply by one-fourth and of sanitation by one-tenth. <i>(It is important that the definitions pertaining to access, adequate amount, safe and sanitary facilities be reported on for the indicators.)</i></p>	<p>Same <i>(but expressed in absolute numbers and as a percentage of the total population).</i></p>

Year 2000 Goal	Year 2000 Goal Indicators	1995 Mid-Decade Goal (MDG)	MDG Process Indicators for HQ, RO and CO Use
Reduction by one-third in the deaths due to acute respiratory infections in children under five years.	<p>Annual number of under-five deaths due to acute respiratory infections.</p> <p>Proportion of pneumonia cases seen at health facilities which receive standard case management.</p>	<p><b><i>For countries that are implementing ARI:</i></b>  Strengthening health facilities capability for case management of pneumonia</p>	<p>Proportion of children under five years of age, with an acute respiratory infection needing assessment, who are taken to an appropriate health provider.</p> <p>Proportion of all health facilities that have a regular supply of free or affordable antibiotics and a trained worker and are thus able to give correct pneumonia case management.</p>

## CHAPTER II

# DECISIONS TO MAKE BEFORE STARTING

*This chapter is written for programme directors, their national counterparts, technical resource persons and survey coordinators. It will help you to:*

- ✓ Identify and contact potential resource persons.
- ✓ Decide on the level of aggregation for the estimates.
- ✓ Find out how long the survey will take.
- ✓ Estimate how much it will cost.

As soon as it is decided that a survey is necessary for measuring the Mid-Decade Goals (MDGs), some important decisions must be made.

- **First**, you must identify a survey coordinator and a senior resource person who will be in charge of the design, implementation and analysis of the data.
- **Second**, you must decide whether a nationally representative sample is sufficient or whether there is also a need for subnational estimates—for example, at the state or province level.
- **Third**, a timetable must be established for the study. Early planning is crucial because the full survey cycle, from inception to publication of results, may take several months.
- **Fourth** and finally, it is necessary to calculate how much the survey will cost.

### IDENTIFYING AND CONTACTING POTENTIAL RESOURCE PERSONS

A **survey coordinator** must be identified in collaboration with government partners. This person will preferably be a professional in a governmental institution who is able to dedicate him/herself full-time for the duration of the project (about six months). The coordinator may also be an independent consultant, but should have the respect of national counterparts. Previous experience in survey implementation and a firm grounding in scientific methodology (e.g., a postgraduate degree in public health or a related field) are also required. This person will make sure that the entire process runs smoothly and that the basic protocols are followed, including carrying

---

---

**Do not attempt to do a survey unless you can identify a full-time survey coordinator. There is no substitute for this important person's role.**

---

---

out the first stages of sampling, selecting and training field workers, supervising overall field activities and processing data. The survey coordinator will also be responsible for obtaining the assistance of the resource person(s) (see below), as well as the cooperation of other government agencies at different levels.

A **senior resource person or institution** should be identified to collaborate closely with the survey coordinator, especially at the survey design stages. This person will have ample technical expertise in survey design, implementation and analysis. A list of such persons will be available from Regional Offices and from the Planning and Coordination Office in New York.

## **DECIDING ON THE LEVEL OF AGGREGATION OF THE INDICATORS**

Sample size depends on whether indicators will be measured at the national or subnational level. Countries are expected to fill major gaps in basic indicators so that they may report on MDGs at the national level. A typical sample for studying **national-level indicators** will be around 4,500 households, in at least 100 clusters spread across the country (see chapter 4 for a full discussion of sample size requirements). It will allow estimation of coverage goals with a margin of error of 5 percentage points or less. This will also usually be adequate for providing baseline estimates to compare with data from a repeat survey in several years time, in order to measure changes. Gender-specific estimates will have a margin of error of about 7 percentage points.

Some countries may also wish to have **subnational estimates** for planning and evaluation purposes. For example, separate estimates for the main provinces or special populations such as slum dwellers may be desirable. Simulations show that samples of 300–500 households per subnational group are sufficient for most MDG indicators (see chapter 4). Subgroup estimates, however, will have wider margins of error (10 percentage points) and will be based on broader age groups than required for reporting on goals at the national level. Alternatively, a subgroup could be oversampled, resulting in an increase in the overall number of households to be surveyed. The potential gains from these options need to be carefully weighed against the additional costs in both time and money that they will inevitably entail.

## **ESTIMATING HOW LONG THE SURVEY WILL TAKE**

Due to the pressing need for reporting, survey planning activities should start as soon as possible. Good planning means specifying clearly at the outset what you need to learn from the survey and how you will use that information. Each step should then be tightly orchestrated around those needs, with the report format and plans for dissemination set out in advance. Experience has shown that without such planning, questionnaire length can get out of control with the inclusion of poorly conceived questions that lack a clear purpose. More important, unless the final phase of the survey



is planned in detail before the outset, data processing, data analysis and report writing will inevitably be delayed, often to the point of rendering the results obsolete by the time they finally appear.

The model timetable presented in Table 2.1 shows the minimum estimated time for completing the full survey cycle for a sample of 150 clusters of 30 households each. This table is provided for general guidance only, since local characteristics may greatly affect the duration of the study. Important factors include the geography of the country, road conditions and previous experience of the survey team, among others.

**Table 2.1. Typical timetable for a national survey covering 4,500 households**

Tasks	Weeks									
	2	4	6	8	10	12	14	16	18	20
Identify resource person and coordinator, plan survey.	■	■								
Carry out sampling, adapt/pre-test questionnaire and training materials.		■	■							
Complete logistic arrangements.			■	■						
Select and train interviewers.					■					
Conduct pilot study and collect data.						■	■	■		
Enter and clean data.						■	■	■		
Complete data processing.							■	■	■	
Prepare report.									■	■

*Assumptions:* 150 clusters of 30 households; one cluster a day with two interviewers. This equals 300 interviewer-days of data collection, or 4 weeks for 20 interviewers allowing travel time and a safety margin.

*A very important issue which may affect the timing of your survey is seasonality. Field work may not be feasible during the rainy season due to poor road conditions, or during special religious holidays such as Ramadam. Also, some indicators may show seasonal variations, such as the prevalence of malnutrition. These influences must be taken into account, particularly if the survey results will be compared with earlier findings.*

## CALCULATING HOW MUCH THE SURVEY WILL COST

Costs will vary widely from country to country. This variability depends not only on currency and labor costs, but also on the degree to which one can reduce costs by using existing facilities. For

example, you can achieve important savings by using government personnel for interviewing, employing public or government-provided transportation, obtaining free accommodations and meals for the survey team from local institutions, and so forth.

Table 2.2 contains a comprehensive checklist of expenditure items. As discussed above, many of the items may not be applicable for particular countries. When applicable, the expenditure is described for a survey with five teams—each team consisting of four interviewers, a supervisor and a driver—spending four weeks in the field. An additional week for training is also budgeted for supervisors and interviewers. Their accommodation costs cover the field work as well as the training period because it is recommended that they be brought to a central location for all training.

**Table 2.2. Common survey budget items and approximate estimates for a survey of 4,500 households**

Budget item	Basis for calculation
<b>Personnel (salaries plus indirect costs)</b>	
Consultants	variable
Field supervisors	5 supervisors x 5 weeks
Interviewers	20 interviewers x 5 weeks
Drivers	5 drivers x 4 weeks
Translators	variable
Local guides	5 guides x 4 weeks
Data entry clerks	2 clerks x 5 weeks
Computer programmers	1 programmer x 8 weeks
<b>Transportation</b>	
Vehicle rental	5 cars x 4 weeks
Public transportation allowance (urban areas)	variable
Fuel	provision for 5 cars x 4 weeks
Eventual costs (repairs, ferries, etc.)	variable
Consultant's visit	variable
<b>Per diems (room and board)</b>	
Field supervisors	5 supervisors x 5 weeks
Interviewers	20 interviewers x 5 weeks
Drivers	5 drivers x 4 weeks
Translators	variable
Local guides (meal allowance)	variable
<b>Consumables</b>	
Stationery (paper, pencils, pens, etc.)	variable
Identification cards	variable
Envelopes for filing	120 envelopes
Computing supplies (paper, diskettes, ribbons, cartridges)	variable

*(table continues on next page)*

---

**(Table 2.2, continued)**

---

**Other Costs**

Questionnaire and form printing .....	6,000 sets
Photocopies of maps, listings and instruction manuals .....	100 maps, etc.
Anthropometric equipment (weighing scales, length meters, etc.) .....	12 sets
Communications (phone, fax, postage, etc.) .....	variable
Report writing and printing .....	variable

---

## CHAPTER III

# DESIGNING THE QUESTIONNAIRE

*This chapter is for survey coordinators. It will enable you to:*

- ✓ Decide which indicators will be measured with the survey.
- ✓ Decide what information you need to collect.
- ✓ Design a good questionnaire.
- ✓ Decide how to ask the questions to obtain the information.
- ✓ Decide who the respondents will be.

### WHICH GOALS CAN BE MONITORED WITH A SURVEY?

Seven Mid-Decade Goals can be monitored with surveys.<sup>1</sup> Each goal has a quantified target to be reached. Box 3.1 summarises these goals.

A set of health indicators, like those defined to measure the Mid-Decade Goals, specifies the minimum information that should be available from the health information system. Full information on the indicators of the Mid-Decade Goals and guidelines for measuring them are found in *Technical Guidelines for Monitoring Progress Toward Mid-Decade Goals*,<sup>2</sup> sent to every UNICEF country office in March 1994.

Many of these indicators rely on routine data collection or collecting special data from health services, but a number of them can be measured in a survey. For some indicators, we must rely entirely upon surveys to provide information representing the whole population. That is why surveys are an integral part of the health information system.

Special modules of questions have been devised to obtain the information needed to estimate indicators of these goals. The modules contain the minimum number of questions needed to estimate these indicators. The Model Questionnaire found at the end of this chapter can be used to obtain survey data on indicators for seven of the Mid-Decade Goals, and for several indicators of two World Summit for Children (Year 2000) goals (see below).

---

<sup>1</sup>*Reporting on Progress towards the Mid-Decade Goals*, UNICEF Executive Directive CF/EXD/1994-001 (10 March 1994).

<sup>2</sup>Programme Instructions, UNICEF CF/PROG/IC/94-003 (15 March 1994).

**Box 3.1****MID-DECADE GOALS THAT CAN BE MONITORED IN A SURVEY**

- Goal 1** ➔ Elevation of immunization coverage of six antigens of the Expanded Programme on Immunization to 80 per cent or more in *all* countries.
- Goal 5** ➔ Virtual elimination of vitamin A deficiency. (At least 80 per cent of all children under 24 months of age in areas with vitamin A deficiency receive adequate vitamin A.)
- Goal 6** ➔ Universal salt iodization in IDD-affected countries.
- Goal 7** ➔ Achievement of 80 per cent usage of ORT (increased fluids) and continued feeding as part of the programme to control diarrhoeal diseases.
- Goal 11** ➔ Reduction of 1990 levels of severe and moderate malnutrition by one-fifth or more.
- Goal 12** ➔ Strengthen basic education so as to achieve reduction by one-third the gap between (a) primary school enrollment and retention rates in 1990 and universal enrollment and retention in primary education of at least 80 per cent of the school-age children, and (b) primary school enrollment and retention rates of boys and girls in 1990.
- Goal 13** ➔ Increased water supply and sanitation so as to narrow the gap between the 1990 levels and universal access by the year 2000 of water supply by one-fourth and of sanitation by one-tenth.

**WHAT ARE THE INDICATORS OF THESE GOALS AND  
WHAT INFORMATION IS REQUIRED TO MEASURE THEM?**

The information necessary to measure the agreed-upon indicators that can be monitored with a survey is described below. Some indicators are more difficult to measure than others, but the lack of available data has pinpointed the urgent need to measure proxy indicators for these goals. The Model Questionnaire has been developed so that countries can choose to include only the modules for goal indicators they need to include in a survey.

---

**To develop your questionnaire, choose only the modules for indicators you need to monitor with a survey.**

---

The Model Questionnaire has been developed so that countries can choose to include only the modules for goal indicators they need to include in a survey.

**EXAMPLE:**

In Zambia, the Breastfeeding, Immunization and Salt Iodization modules were chosen to add on to a crop forecast survey to be carried out twice every year.

Some countries will have data to measure some indicators from routine sources and will not need to include all the modules in a survey. The more modules that are included, the more complex the survey becomes.

**Goal 1:** Elevation of immunization coverage of six antigens of the Expanded Programme on Immunization to 80 per cent or more in *all* countries.

*Indicator (1.1–1.4):* proportion of children receiving each immunization (DPT3, measles, OPV3, BCG) before first birthday.

*Indicator (1.6):* proportion of children protected against neonatal tetanus through immunization of their mother.

Many countries already have information from routine data collection to monitor indicators 1.1, 1.2, 1.3 and 1.4. If you choose the question modules for these indicators of immunization coverage, you will find that several options for collecting these data are given. One option, given in the standard Immunization Module found in the Modules for Children Under Five Years of Age, is to record all vaccinations and the dates they were given, as in the EPI survey. Another option, used in the Alternative Immunization Module, is to record only whether the dose was given, from either the card or the mother's verbal report, and gives alternative probing questions to use when no card is available. Choose the option that best serves your country's monitoring needs, and suits the demands of your survey.

*Indicator 1.6* is more difficult to monitor with a rapid survey than child immunization status. In many countries validation of a mother's immunization at the time of a child's birth with dates from a card may be difficult because many mothers will not hold cards for this immunization. When few mothers hold cards, asking questions about immunizations during the last pregnancy and probing further back in time about previous doses of tetanus toxoid received may produce the best information. The Tetanus Toxoid Module is found in the Modules for Mothers (beginning on page 3.28 [Q4]).

**Goal 5:** Virtual elimination of vitamin A deficiency. (At least 80 per cent of all children under 24 months of age in areas with vitamin A deficiency receive adequate vitamin A.)

*Indicator (5.1):* proportion of children under 2 years of age receiving adequate vitamin A (in known deficient areas only).

In the absence of information on vitamin A deficiency, it is very likely that areas with a high prevalence of child protein-energy malnutrition, high infant and young child mortality rates, known seasonal food shortages and periodic droughts, and areas with a high prevalence of diarrhoea and/or measles will also be areas where vitamin A deficiency in young children is prevalent. On the basis of this information, it should be possible to map major regions of the country according to the risk of vitamin A deficiency. If no programme is in operation in these defined "at risk" areas, and no further information is available, the goal will not be met.

If a programme to eliminate Vitamin A deficiency is under way, then proxy indicators of programme progress can be measured, but a multiple-indicator survey is *not* a suitable vehicle for assessing Vitamin A deficiency.<sup>3</sup>

As agreed by WHO and UNICEF,<sup>4</sup> the operational definition of Indicator 5.1 will vary according to the programme strategy undertaken by a country, as follows:

**For countries/areas with a Supplementation Programme:** proportion of children reaching 24 months of age who had received at least 400,000 IU of vitamin A from supplements given at appropriate intervals according to established policy, and/or the proportion of mothers given high-dose supplements as a proportion of the number of children immunized against BCG.

OR

**For countries/areas with a Food Fortification Programme:** proportion of 12–24-month-old children receiving at least 1,100 micrograms retinol equivalents per week from fortified foods.

OR

**For countries/areas with a Dietary Diversification/Education Programme:** proportion of 12–24-month-old children eating a diet providing at least 2,200 micrograms retinol equivalents per week.

The indicators described above cannot be measured by asking simple questions incorporated into a multi-purpose monitoring survey. More detailed dietary assessment surveys are needed to assess the adequacy of vitamin A intake, and to assess the proportion of the population receiving an

---

<sup>3</sup>See *Methodologies for Monitoring and Evaluating Vitamin A Deficiency Intervention Programmes* (Washington, D.C.: International Vitamin A Consultative Group Secretariat, 1994).

<sup>4</sup>These indicators are explained in more detail in *Technical Guidelines for Monitoring Mid-Decade Goals*, CF/PROG/IC/94-003 (15 March 1994).

adequate amount of vitamin A. However, a monitoring survey will be able to provide useful information which can help countries decide whether they are likely to be on track in reaching the goals, and whether in-depth surveys to look more closely at quantitative indicators of vitamin A intake are needed.

The questions in the Vitamin A Module of the Modules for Children Under Five Years of Age are designed only to measure the extent to which the programme under way in the given country or area of a country is reaching those to whom it is targeted. The modules are *not* designed to measure the extent of vitamin A deficiency or the adequacy of vitamin A intake. They are not designed to measure the frequency with which supplements, vitamin A-rich foods or foods fortified with vitamin A are consumed.

The question modules *are* designed to measure the proportion of the population (either families or children) reached by the programme:

- through measuring the proportion of children who have received a vitamin A supplement within the prescribed time period, or
- the proportion of households with a fortified food product, which is given to children in those households, or
- the proportion of mothers who have heard the programme message and put the message into practice (by assessing the proportion of children who eat the target foods).

The choice of questions is determined by the strategy undertaken. The modules give examples of questions that can be used as models for programme-specific questions. You will need to tailor the wording of these questions to fit your country's vitamin A intervention programme. In some cases, this will mean compiling a list of vitamin A-rich foods available in particular areas during different seasons of the year.

---

---

**The vitamin A modules (A, B or C) are designed to monitor programme coverage. Therefore, only those countries with a defined strategy for vitamin A deficiency and an active programme should use the questions.**

---

---

**Goal 6:** Universal salt iodization in IDD-affected countries.

*Indicator (6.1):* proportion of households consuming adequately iodized salt according to agreed criteria (i) in the whole country, and (ii) in areas known to be at high risk of IDD.

Testing kits<sup>5</sup> for salt iodization are used to test the salt used in the household. The questionnaire items are found in the Salt Iodization Module of the Household Questionnaire.

---

<sup>5</sup>See chapter 5 of this handbook for ordering information.



**Goal 7:** Achievement of 80 per cent usage of ORT (increased fluids) and continued feeding as part of the programme to control diarrhoeal diseases.

In 1993, the definition of oral rehydration therapy (ORT) changed. ORT is now defined as increased intake of acceptable fluids, including plain fluids found in the home and oral rehydration solution. Both the pre-1993 definition of ORT use and the current definition of ORT (increased fluids) and continued feeding can be measured in the survey, as follows:

*Indicator (7.1):* proportion of all diarrhoea episodes in children under five years of age treated with oral rehydration salts (ORS) and/or recommended home fluids. Use of ORT (pre-1993 definition).

*Indicator (7.2):* proportion of diarrhoea episodes in under-fives treated with ORT (increased fluids) and continued feeding. Use of ORT (increased fluids) and continued feeding.

Suggested questions and response categories, agreed upon by WHO and UNICEF, are given in the Diarrhoea Module of the Modules for Children Under Five Years of Age. Individual country programmes will need to decide how to tailor response categories to the country setting for questions about treatments and fluids given for a diarrhoeal episode, and decide on the codes for "correct" responses.

**Goal 11:** Reduction of 1990 levels of severe and moderate malnutrition by one-fifth or more.

*Indicator(11.1,2):* proportion of under-fives who fall below minus 2 (minus 3) SD from median weight for age.

(OPTIONAL) *Indicator (11.3, 4):* proportion of under-fives who fall below minus 2 (minus 3) SD from median height for age.

Obtaining data for these indicators will require that children under age five are weighed (and, optionally, measured). The Anthropometry Module in the Modules for Children Under Five Years of Age contains items to record weights of under-fives in each household and (optionally) height or length. More information about the procedures to measure weight and height are found in Appendix 2.

**Goal 12:** Strengthen Basic Education so as to achieve reduction by one-third of the gap between (1) primary school enrollment and retention rates in 1990 and universal enrollment and retention in primary education of at least 80 per cent of the school-age children, and (2) primary school enrollment and retention rates of boys and girls in 1990.

*Indicator (12.1):* proportion of children entering first grade of primary school who eventually reach grade 5.

*Indicator (12.2):* number of children enrolled in primary school who belong in the relevant age group (e.g., of primary-school age, as defined in country), expressed as a percentage of the total number in that age group (net enrollment).

*Indicator (12.3):* proportion of children of primary-school entry age who enter grade 1 at that age.

*Indicator (12.4):* percentage of children enrolled in primary school of total number of primary-school-age children (gross enrollment).

Questions to measure these indicators are found in the Education Module, administered to mothers or caretakers of children over the usual age of school entry who have not yet reached age 15.

**Goal 13:** Increased water supply and sanitation so as to narrow the gap between the 1990 levels and universal access by the year 2000 of water supply by one-fourth and of sanitation by one-tenth.

Questions to measure indicators of this goal are found in the Water and Sanitation Module of the Household Questionnaire. These questions should be asked in every household in the survey sample, whether or not eligible children and their mothers reside in the household.

*Indicator (13.1):* population with access to an adequate amount of safe drinking water located within a convenient distance from the user's dwelling.

A consistent definition of what is considered "safe and convenient" should be established before the first survey takes place. Type of water source is used to indicate "safe" supply. Possible categories for type of water source and distance from water source are shown in model questions 1 and 2. In some countries it may be appropriate to categorize "convenient distance" in terms of the time it takes to travel to the source, collect water and return; in some, distance can be estimated and categorized. It is suggested that countries not attempt to quantify "adequate amount" of water through this survey.

*Indicator (13.2):* population with access to a sanitary facility for human excreta disposal in the dwelling or within a convenient distance from the user's dwelling.

Definitions of what is considered a "sanitary facility" and "convenient distance" should be decided, and appropriate country-specific categories, covering likely responses, should be used. Suggested categories appear in model questions 3 and 4.

Some countries may wish to add further questions to this module. These may include questions to provide information on reliability of the current water source and who is responsible for maintaining operation of the source. Exercise caution when adding questions to a multi-purpose survey, so that field staff and respondents are not overburdened by the demands of the survey instrument.

 **EXAMPLE:**

A country may want to determine the reliability of the water source named by the household.

Ask: "Can you get water from this source everyday, most of the time, not very often or almost never?"

In addition to indicators of the seven Mid-Decade Goals, it is possible to measure indicators of the **World Summit For Children (WSC) Goal 16**: *empowerment of all women to breastfeed their children exclusively for 4–6 months, and to continue breastfeeding, with complementary food, well into the second year*. The indicators to monitor progress at mid-decade (**Mid-Decade Goal 8, Breastfeeding and Baby-Friendly Hospital Initiative**) are measured using data from health facilities, and cannot be measured with a household survey.

The following indicators have been devised for those countries that want to monitor breastfeeding using a survey.<sup>6</sup> The Breastfeeding Module is found in the Modules for Children Under Five Years of Age.

*Indicator*—Exclusive breastfeeding: proportion of infants less than 4 months of age exclusively breastfed.

*Indicator*—Timely complementary feeding: the proportion of infants 6–9 months of age receiving breast milk *and* complementary foods.

*Indicator*—Continued breastfeeding at 2 years: the proportion of children 20–23 months who are breastfeeding.

*Indicator*—Continued breastfeeding at 1 year: proportion of children 12–15 months who are breastfeeding.

Definitions of "complementary foods" can be made country-specific, if country programmes promote certain complementary foods as appropriate or adequate.

(OPTIONAL) *Indicator*—Bottle-feeding rate: proportion of infants less than 12 months of age receiving any food or drink from a bottle (current status).

---

<sup>6</sup>See *Indicators for Assessing Breastfeeding Practices*, WHO/CDD/SER/91.14 (June 1991).

(OPTIONAL) *Indicator*—Ever breastfed rate: proportion of infants less than 12 months of age who were ever breastfed.

The "current status" approach is adopted, to assess current breastfeeding practices at the time of the survey. However, numbers of children encountered in the age ranges of interest are likely to be quite small: 0–3-month-olds, 6–9-month-olds and 20–23-month-olds. These questions may be included in the survey, but to stay within feasible sample sizes, the precision with which they are measured may need to be less demanding than that for other indicators. There are well-known problems obtaining good data about breastfeeding duration when a mother has already stopped breastfeeding a child, but one question to be used in countries where breastfeeding durations are known to be very short (i.e., under 6 months) is also given.

One other indicator, maternal knowledge of referral in acute respiratory infection, can be monitored with a survey.<sup>7</sup> This is a key indicator that some countries may wish to use to monitor programme progress toward **WSC Goal 24**: *reduction by one-third in the deaths due to acute respiratory infections in children under five years*.

**Strategy for Control of Acute Respiratory Infections (CARI)**: Early recognition and prompt referral of pneumonia by parents and family.

(OPTIONAL) *Indicator*: proportion of mothers (caretakers) of children under five years who know the signs (fast/difficult breathing) which indicate that a child with a cough and/or cold must be taken to an appropriate health provider.

The CARI questionnaire module is found in the Modules for Mothers.

The **World Summit for Children Goal 1**: *Between 1990 and the year 2000, reduction of infant and under-five mortality rate by one-third or to 50 or 70 per 1,000 live births, respectively, whichever is less*, can also be monitored in a survey. A separate module of questions designed to monitor childhood mortality is also included in the Model Questionnaire.

*Indicator*—Infant mortality rate: the annual number of deaths of infants under 1 year of age per 1,000 births.

*Indicator*—Under-five mortality: the probability of children dying between birth and their fifth birthday, expressed per 1,000 children born alive.

---

<sup>7</sup>UNICEF, *Guidelines for Monitoring the Mid-Decade Goals: ORT, CDD, CARI* (New York: UNICEF, January 1994).

Complete information on how to administer these questionnaire modules and how to analyse the data is found in another UNICEF publication, *Measuring Childhood Mortality: A Guide for Simple Surveys*.<sup>8</sup>

### WHAT ARE THE REQUIREMENTS OF A GOOD QUESTIONNAIRE?

A survey to measure indicators of the Mid-Decade Goals will have multiple aims, which means that the data collection process immediately becomes more complex than an EPI or other single-purpose survey.

Be clear about your aims—make sure the questionnaire for your survey is relevant to your purposes. Don't make the survey too long by including many unnecessary questions. It is always tempting to add more and more questions, expanding the survey unnecessarily. Resist this temptation, and collect only the

minimum information you need. Otherwise you run the risk of overloading your field workers, and demanding too much of your respondents. Each question should have an explicit rationale for inclusion. Why is it included and what will be done with the information after it is collected?

The main aim of a good survey instrument is to minimize the amount of error that can occur when measuring whatever it is you want to measure. This means that a good questionnaire can be used by interviewers to obtain answers that are both *reliable* and *valid*. By reliable, we mean that no matter who asks the question, and no matter where and when it is asked, the same respondent would give the same answer. In a good questionnaire, the same question is asked in the same way by different interviewers—and differences between interviewers will be kept to a minimum. By valid, we mean that the question elicits a response that measures whatever it is

you are interested in measuring and that the answer given to the question is true and accurate. A good questionnaire should enable you to obtain valid measures of the things you set out to measure, by helping to ensure that the respondent understands what information is being sought.

---

---

**Be clear about the data required to meet survey aims. A questionnaire should always contain the *minimum* number of questions needed to obtain the information.**

---

---

---

---

**Collect reliable and valid information.**

---

---

---

<sup>8</sup>P. H. David, L. Bisharat and A. G. Hill, *Measuring Childhood Mortality: A Guide for Simple Surveys* (Amman: UNICEF Middle East and North Africa Regional Office, 1990).

**Box 3.2**  
**SOURCES OF ERROR IN SURVEYS**

Data from retrospective surveys may contain errors for many reasons. These errors can be grouped under two main headings:

**Sampling error** arises by chance, because the survey asks questions of a sample of respondents rather than the whole population. Errors can also arise because your sample is not representative of the entire population. You can avoid these kinds of errors by ensuring that your sampling frame is adequate and your sample size is large enough to enable your measurements to be precise. In chapter 4, we discuss the ways to avoid sampling and coverage errors.

**Measurement error** results from imperfectly measuring what you set out to measure. This kind of error is usually even more serious than sampling error because it cannot be corrected, and sometimes not even detected, when it happens. One important way to avoid measurement error is to ensure that your survey instrument is carefully designed. Chapter 3 tells you how to design your questionnaire. Another way to avoid measurement error is to ensure that the interviews are conducted by well-trained and supervised interviewers. Chapter 5 discusses this aspect of the survey in detail.

These considerations are especially important for monitoring surveys because the purpose of a survey is to measure trends in the indicators over time, and to compare indicators internationally.

---

---

**In monitoring surveys, the questions must  
be asked in the same way each time a  
survey is conducted.**

---

---

The survey instrument is the key to these two important criteria, but there are alternative ways to structure it, and each has arguments in its favour. You can use a *verbatim* questionnaire, in which the questions appear on the questionnaire exactly as they are to be asked, translated so that both interviewer and respondent have the same understanding about what is being asked. It is important that no ambiguous words or phrases are used in a questionnaire, to guard against misinterpretation. A verbatim questionnaire structures the interview and reduces variability between interviewers in the way they pose the questions. That is especially important in situations where responses can change a great deal, depending on the exact wording of the question.

The other alternative is a simple listing of the information needed, such as forms used for medical records. This saves time and space, and allows a great deal of flexibility. These forms usually contain a list of physical signs and symptoms to be ticked by the doctor or health worker. This kind of summary form is often used in EPI surveys.

Questions posed in multiple-indicator monitoring surveys, however, may be more general. They may ask for responses that can be conditioned by the way the question is asked. While the summary form may save on paper, in a more complex survey to measure

---



---

**The *questionnaire* format of the Model Questionnaire Modules, with each question appearing verbatim on the form, is preferable to a summary form.**

---



---

multiple indicators it opens the way to many errors. The Model Questionnaire Modules to measure the Mid-Decade Goals are designed to be adapted for use all over the world. The verbatim questionnaire is most likely to ensure that results of surveys to measure the Mid-Decade Goals are comparable with results from other surveys conducted in other countries and with results from surveys that were repeated in the same country every few years.

 **EXAMPLE:**

When you ask a mother, "Would you take your child to a clinic if he had difficulty breathing?" she is likely to answer "yes" because she knows that is the "correct" or expected answer. You may get a very different answer to the question, "When your child has a cough or cold, what signs would prompt you to take him/her to a clinic?"

The EPI form gives only a space for the answer the mother gives, but does not contain the question as the interviewer should ask it:

Immunization Card	Yes/No
BCG	Date/+/0
Source	_____

Instructions on the form give only the key for responses given, not the questions to ask when a child has no immunization card. *You should write out the question just as you want the interviewer to ask it.*

The other important aim of a good survey questionnaire is to obtain the necessary information quickly and easily. This means that, as noted earlier, it should contain the

---



---

**A good questionnaire is accurate.**

---



---

minimum number of questions needed to obtain the required data, and both the interviewer and the respondents should easily understand these questions. The wording and sequence of questions are designed to motivate respondents and help them recall difficult information. In this way, the survey should be manageable by country staff, economical in both time and money, and will intrude as little as possible on the activities and privacy of families who are interviewed.

 **EXAMPLE:**

Do not confuse the respondent by asking two questions at once, such as "How many children over the age of five are there and do any of them attend school?" Instead, ask two separate questions: "Can you give me the names and ages of each child in the household?" Then, for each child over five, ask, "Is [NAME] currently at school?"

Even when you have constructed a good questionnaire, it is not a guarantee that interviewers will stick to the correct interpretation of the questions. Good training in the use of the questionnaire is also essential. Instructions for administering the questionnaire are found in Appendix 1. Advice on how to select and train interviewers is found in chapter 5.

A guide to help you calculate each indicator from the data obtained using a Model Questionnaire is found in chapter 7, and a computer program to analyse the data and to calculate each indicator accompanies this

---

---

**Your interviewers must learn how to ask questions properly. Be sure you use the interviewer guide to train your field staff.**

---

---

handbook (see Appendix 3). Sample sizes permitting, indicators can be tabulated for the whole country, and for each region, urban/rural residence, sex, age and socioeconomic status.

Several suggested questions to assess the socioeconomic status of the household are found on page 1 of the Model Questionnaire. If you want to examine within-country socioeconomic differentials for the indicators, you can adapt or change these questions to make them appropriate for use in your country.

 **EXAMPLE:**

Some countries may find that a question about education of the head of household is a good indicator of a family's wealth.

Ask: "Have you ever been to school?" and "For how many years did you attend school?"

Others know that information on the condition of the dwelling is a better proxy for a family's socioeconomic status, and families with a dirt floor are the poorest families.

Ask: "What is the dwelling floor made of?"



**Box 3.3****POINTS TO REMEMBER WHEN DESIGNING YOUR QUESTIONNAIRE**

- Use simple, direct language.
- Make questions clear and specific.
- Do not combine questions—ask them one at a time.
- When questions do not apply to all respondents, include SKIP instructions to filter out respondents.
- Make sure questions do not lead to “expected” answers; use wording which allows each response option to be given equal consideration.
- Put difficult or sensitive questions at the end of the questionnaire, if possible.
- Respect the respondent's right to privacy.

**Decide How to Ask the Questions*****Adapting and Translating Questionnaires***

No changes should be made to the wording of the questions; they have been carefully designed to measure each indicator. However, the **response** categories for some questions may need to be adjusted to the requirements of different country settings. These response categories may be changed, but you should follow some guidelines for making these changes. If you make changes to the response categories, make sure that the options listed are mutually exclusive (one response does not fit more than one category), and that the categories available exhaust all possible responses. Be sure to allow for a residual category for responses which do not fit into other categories. This is so that interviewers will have a place to record all answers given, a code will be given to each answer and no blank spaces will be left when the questionnaire is completed.

** EXAMPLE:**

You are questioning a mother about her baby's diarrhoea.

Ask: "During Fatima's diarrhoea, did she drink much less, about the same or more than usual?"

Categories for the interviewer to record the mother's response are: (1) much less or none, (2) about the same (or somewhat less), (3) more, (4) don't know.

It will usually be important to compare data obtained from one survey with another, and so

it is important to keep the questions and response categories the same every time the questionnaire is used. If the data collection method differs from one survey to the next, the data may no longer be valid and comparable. Make sure the response categories you choose are detailed enough to allow comparison with other data sources. The new response categories should be pre-tested.

- ☞ The number of response options should be as small as possible and contain a category for responses that do not fit any specific option.

### ***Planning the Data Management and the Analysis***

Any change to the format of the Model Questionnaire will require that changes be made to the data entry and calculation programs provided with this handbook.

It is important to plan how the data will be entered into a computer and how they will be tabulated early in the planning process. You should construct dummy tables, based on the response categories you have chosen, to ensure that you can tabulate each indicator from your questionnaire.

**If you decide to change or add questions to the modules, you will need an excellent programmer familiar with EPI INFO to make the necessary changes to these computer programs.**

#### **📎 EXAMPLE:**

Construct a dummy table indicating mothers of under-twos who have heard the programme message about vitamin A-rich foods and the number who mention the food promoted in the message.

	<b>Mothers heard message</b>		
<b>Mentions green leafy vegetable</b>	Yes	No	Total
Yes			
No			
Total			

### ***Translating the Question Modules***

These question modules need to be translated into the expected respondents' local language *before* the survey begins. Translation should never be left to the interviewer; small differences of interpretation can destroy the reliability and validity of your data.

In a separate operation, *another translator* should then translate the new questions back into the original language again (without reference to the original model). This new English translation should match the original English version. Any ambiguous words or phrases should be discussed, and the correct translation into the local language should be decided.

### ***Pretesting the Questionnaire***

The translated questionnaire should be *pre-tested* in the community, with respondents similar to the respondents in the survey sample. You will find more information about doing a pre-test in chapter 5. This pre-test should identify problem areas, misinterpretations and cultural objections to the questions.

The pre-test of your questionnaire is especially important—it is easy to get it wrong the first time. A pre-test can provide a great deal of information to use in designing the final questionnaire and for planning other aspects of the survey process.

#### **Box 3.4**

##### **WHAT A PRE-TEST CAN TELL YOU**

- ◆ Are respondents willing to answer the questions in the form you propose to use?
- ◆ Are any of the questions particularly difficult or sensitive? Extra training can focus on these questions.
- ◆ Are the questions misinterpreted by the respondents? Are any of the words ambiguous or difficult to understand? The pre-test should point to where changes in wording or improved translation is needed.
- ◆ Does the questionnaire flow smoothly? Can the interviewers follow the instructions easily? Do the interviewers misinterpret the questions?
- ◆ Is there adequate space on the form and are the answers clearly coded? The pre-test should show where the format needs to be improved before the final questionnaire is printed.
- ◆ How long does an interview take? The answer to this question will help you decide how many interviewers are needed and how long the field work will take.

You may find that the response categories for some questions are not sufficient to allow for the range of answers you receive. For example, you may need to add a very popular

children's drink to the list of fluids given to a child during an episode of diarrhoea (See Diarrhoea Module). Or, you may find that there are too many categories for some questions in the suggested Model Questionnaire. For example, there may only be three or four sources of drinking water available to households in your sample (see Water and Sanitation Module). You can shorten the list of response categories, as long as you allow for a residual category and you are sure that your response groups allow you to compare your results with other sources of data.

 **EXAMPLE:**

If you know that many households get water from a cistern to collect rain water, and you want to know more about which ones obtain water in this way, you may make this a separate category.

Ask: "What is the source of drinking water for members of your household?"

1. piped-in dwelling
2. public tap
3. tube well, borehole, protected well or spring
4. unprotected well or spring
5. rainwater collected in cistern
6. pond, river or stream
7. tanker truck or vendor
8. other source

Do not ignore the lessons of the pre-test. But do not change the order or the meaning of the questions. You may need to do more than one pre-test before your questionnaire is satisfactory. Discuss the results with experi-

enced colleagues and with the interviewers, and decide what changes are needed. Make any changes necessary to the instructions to interviewers, to the wording of prompting questions and to the translation. Only then are you ready to reproduce the questionnaire forms.

In summary, the final questionnaire should be the product of careful preparatory work. It should ask only for the information your programme needs. It should be as short as possible, but contain questions that elicit both reliable and valid responses. It should be as easy as possible for interviewers to use, while ensuring that the data obtained are accurate.

---

---

**The results of the pre-test should be incorporated into a final questionnaire.**

---

---

---

---

**Do not rush to print your questionnaire before you have done the pre-test and made the necessary changes.**

---

---

*If the wording, order or layout of the questions is changed, special care should be taken to ensure that all the necessary information to estimate the indicators will be obtained.*

**Box 3.5**  
**QUESTIONNAIRE CHECKLIST**

- ✓ Questions should be clear, use simple language and be as short as possible.
- ✓ The questions should flow clearly and logically from one to the other, and the layout should make it easy to administer the questionnaire.
- ✓ Appropriate codes should be used for all response categories, and a code should be available for a category of answers that do not fit into the other response categories.
- ✓ Instructions to interviewers should be easy to follow and easy to distinguish from the questions to the respondents.
- ✓ Make sure the questionnaire has been translated, back-translated and pretested.
- ✓ Check that the questionnaire provides *all* the information needed to calculate estimates of each indicator.

### **Decide Who the Respondents Will Be**

The model questionnaire to measure the Mid-Decade Goals is designed to be administered to every mother or caretaker of young children in each household drawn for the survey sample. Mothers or, in their absence, the principal caretaker, are asked questions about the education and health of the children in their care. The mother or principal caretaker is most likely to be able to give accurate answers to the questions. Mothers of under-fives are also asked about their own immunizations and about their knowledge of acute respiratory infections. The head of household may answer the questions about household water and sanitation and salt iodization. When the Anthropometry Module is included in the survey, each child under five living in the household must be weighed (and sometimes measured, when height-for-age is a chosen indicator).

You will see that the Modules for Mothers and the Modules for Children Under Five Years of Age are directed to mothers or caretakers of all children under age five. Most of the indicators measured in this survey focus on a specific age range: all under-fives for ORT use and nutritional status, under-twos for vitamin A programme coverage, under-ones for tetanus toxoid coverage, one-year-olds for other immunization indicators and different age groups for breastfeeding. The questions

are asked for all under-fives to enable you to obtain some subnational estimates based on the larger age range (see chapter 4). The calculations of national-level estimates are based on the specific age range for the indicator required for reporting on the Mid-Decade Goals.

A survey that includes mortality measurement must interview *all* women of reproductive age in the households sampled, whether or not they have living children. For further information about designing a survey that includes measuring the level of childhood mortality, consult the UNICEF publication *Measuring Childhood Mortality: A Guide for Simple Surveys*, by P. H. David, L. Bisharat and A. G. Hill (Amman: Middle East and North Africa Regional Office, 1990), available from the Office of Planning and Coordination, New York.

---

---

**The target respondents for a survey that includes childhood mortality measurement differ from those for a survey to monitor indicators of the Mid-Decade Goals.**

---

---

## THE MODEL QUESTIONNAIRE

The model questionnaire for measuring the Mid-Decade Goals asks for the minimum information necessary to measure the chosen indicators. Short, specific questions using simple, everyday words are asked. One question is asked at a time. The structure of the questions, the wording used and the layout and order of the questions have been carefully chosen to ensure that all necessary information to monitor each indicator is collected.

The questionnaire is not intended to fulfill the functions of more elaborate EPI, CDD or other specific programme evaluation surveys. It is intended for use in surveys which aim to measure multiple indicators, for reporting progress toward the Mid-Decade Goals.

### Organization of the Questionnaire

The questionnaire is set up as a series of modules. Country programmes can choose those modules that measure selected indicators they intend to monitor with a household survey.

The rationale for the organization of the questionnaire is:

1. The standard format for the questionnaire begins by collecting a few basic items of information about the household. For countries wishing to examine regional or socioeconomic differentials in the indicators, examples are given of possible questions to use to ascertain socioeconomic status.

2. A list of all mothers and children living in the household is obtained. In the mother's absence, the principal adult caretaker of these children is interviewed. The interviewer starts with the youngest mother, recording her name and age on the form. Then, each child's name, sex and date

of birth should be entered on the form. Special checks are used to make sure the date of birth and age obtained are accurate. The correct age is especially important when the Anthropometry Module is included in the survey. When a child over the age of 15 years is reached, the interviewer is instructed to stop listing, and to go on to other mothers/caretakers and their children who reside in the household.

3. Even in households where no mothers or young children reside, a questionnaire containing questions about water and sanitation and salt iodization is administered. These questions are to be answered only once for every household in the survey, and can be answered by any adult in the household.

4. For each child over the usual school-entry age, questions in the Education Module about school attendance are completed. The mother or caretaker of each eligible child is questioned.

5. The interviewer then moves to the Modules for Mothers, which include the Tetanus Toxoid Module and the CARI Module. These are addressed to all mothers of under-fives in the household. When the Tetanus Toxoid Module and the CARI Module are finished, the instructions take the interviewer to the Modules for Children Under Five Years of Age.

6. For each child under age five, a separate module is filled in. The Modules for Children Under Five Years of Age contain all the chosen health modules. The child's line number and name, as well as household and cluster identifying numbers, are entered on each page.

7. The first module, usually the Diarrhoea Module, is administered. If the child has not had a recent diarrhoea episode, a "skip" instruction takes the interviewer to the next module.

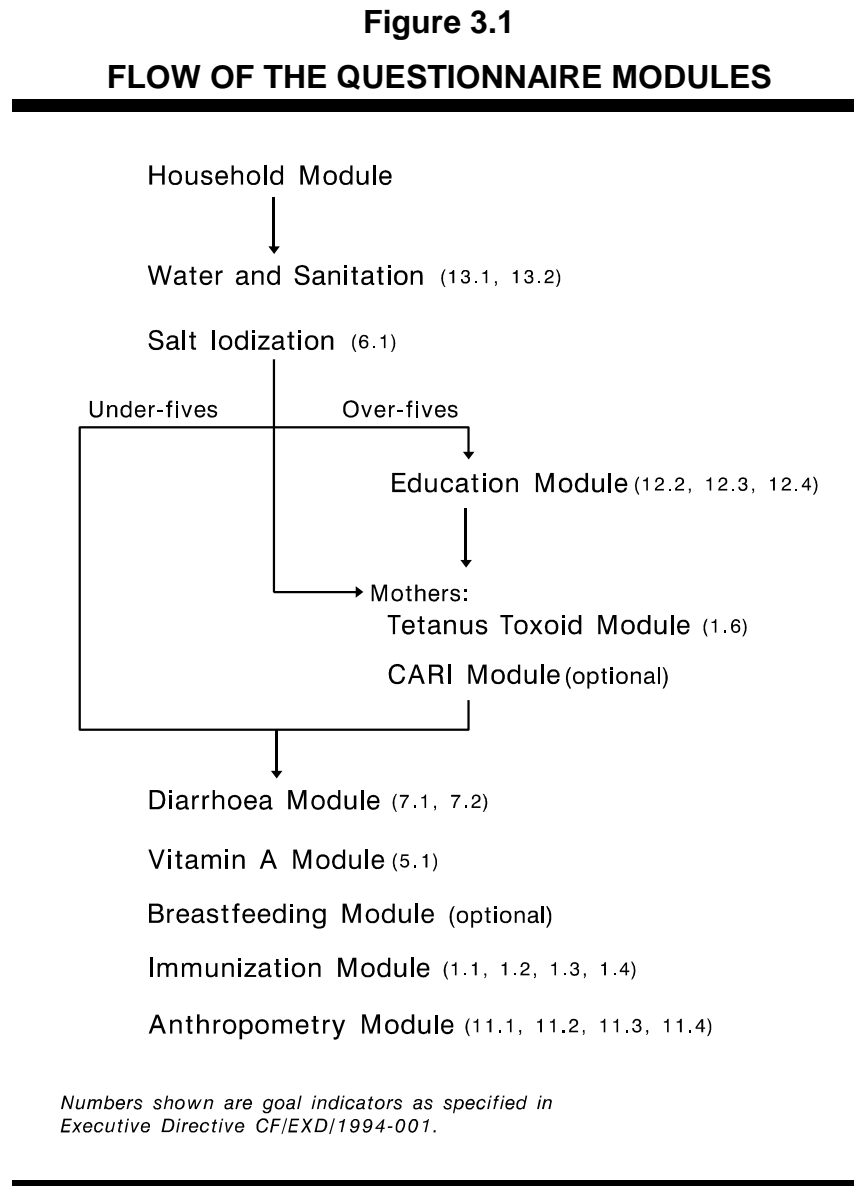
8. The next module is the Vitamin A and/or Breastfeeding Module. Although the indicators to be measured at the national level refer to specific age ranges, data are gathered on all under-fives. This will enable some subnational estimates of the indicators to be made, based on the larger group of children under age five.

9. Following the Vitamin A and Breastfeeding Modules, the interviewer moves to the Immunization Module, where data on immunizations are obtained.

10. Finally, the interview ends with the Anthropometry Module for all children younger than five years of age. An option is to place a notation for presence of BCG scar in the Anthropometry Module, too, since it may require undressing the child. A measurer and assistant should be included in the team used for the anthropometry measurements. If possible, the interviewer makes up the third member of the interviewing team, left free to record the measurements or to go on to the next child.

One set of Modules for Children Under Five Years of Age is filled in for each child under five in the household, and then the team proceeds to the weighing and measuring. The anthropometry is left until last because it is likely to cause the most disruption to the interviewing process.

The flow of the Model Questionnaire is found in Figure 3.1.







## MODEL QUESTIONNAIRE MODULES

The model questionnaires that follow are not intended to be completely self-contained: use the detailed instructions for the interviewer (provided in Appendix 1) in conjunction with the modules. In addition, general instructions are provided in italic text (*instructions*) in the questionnaires themselves, except for the instruction to proceed to the next module, which is capitalized and bold (**GO TO NEXT MODULE** ⇨). Questions that the interviewer will be asking appear in small capital letters (INTERVIEWER QUESTION) in the questionnaires, to distinguish them from the general instructions. Because of space limitations in some of the modules, the response "don't know" is sometimes abbreviated as "DK," and "household" is abbreviated as "HH."

To include mortality measurement in multiple-indicator surveys, use the alternative listing sheet and Mortality Module on page 3.36 (Q12).

**HOUSEHOLD MODULE**

**INTERVIEWER:** *Begin by introducing yourself—for example, WE ARE FROM \_\_\_\_\_ AND WOULD LIKE SOME INFORMATION THAT WILL HELP US IMPROVE THE HEALTH AND WELL-BEING OF CHILDREN. THE QUESTIONS WILL TAKE ONLY A FEW MINUTES.*

**Household Information Panel**

Cluster number:	Household number:	Date of interview (day/month/year): __ / __ / __
Interviewer no.:	Name of head of household:	Call-back necessary? Yes / No Time: __ A.M./P.M.
No. persons in HH usually resident:	Material of dwelling floor: 1 wood/tile 2 planks/concrete 3 dirt/straw 4 other	Number of rooms in dwelling:
Data entry clerk no.:	All forms completed? 1 Yes 0 No	Region 1 2 3 4
	If not, why not? 1 Refusal 2 Not at home 3 HH not found/destroyed 4 other	Urban 1 Rural 2

**INTERVIEWER:** I WOULD LIKE TO ASK ALL MOTHERS OR OTHERS WHO CARE FOR CHILDREN SOME QUESTIONS ABOUT THE HEALTH AND WELL-BEING OF THE CHILDREN IN THIS HOUSEHOLD. *Ask to speak to each mother/caretaker, listing the first mother’s name in line 1-0. (See Appendix 1 for more information.) Ask each woman in turn to list the names and birth dates of the children for whom she is responsible who live in the household, starting with the youngest child, who is listed on line number 1-1. Stop listing when you reach a child over age 15. Go on to the next woman, listing her name and the children for whom she is responsible who are living in the householder, starting with line number 2-0. Repeat the procedure for each mother/caretaker in the household. Add a continuation sheet if there is not enough room on this page. Then ask: ARE THERE ANY OTHER CHILDREN WHO LIVE HERE, EVEN IF THEY ARE NOT AT HOME NOW? (These may include children in school or at work.) If yes, complete listing. Then, ask and record answers to questions as described in the Instructions to Interviewers. Tick here if you use a continuation sheet .*

**Mother and Child Listing Form**

Line no.:	1. Name:	2. Is ____ male or female? 1=male 2=female	3. Date of birth? (dd/mm/yy)	4. Age in years	5. Age 5 or over? <sup>a</sup>
1-0		1 2			1 2 3 4
1-1		1 2			1 2 3 4
		1 2			1 2 3 4
		1 2			1 2 3 4
		1 2			1 2 3 4
		1 2			1 2 3 4

<sup>a</sup>For question 5, 1=yes, 2=no, 3=mother is respondent, 4=caretaker is respondent. Copy line numbers for all children age 5 and over to next page. Copy line numbers for all under-fives to separate Modules for Children Under Five Years of Age, one for each under-five, containing all health modules.

**WATER AND SANITATION MODULE**

Cluster no. \_\_\_ Household no. \_\_\_

Ask the questions in this module once for each household visited. Record the number for only one answer in the space at right. If a respondent gives more than one answer, enter the most usual source/facility:

## 1. WHAT IS THE SOURCE OF DRINKING WATER FOR MEMBERS OF YOUR HOUSEHOLD ?

Piped-in dwelling	1	Unprotected dug well or spring,	
Public tap	2	rainwater	5
Tube well or borehole	3	Pond, river or stream	6
Protected dug well or protected spring	4	Tanker-truck, vendor	7
		Other	9

## 2. HOW FAR IS THIS SOURCE FROM YOUR DWELLING ?

On premises	1	500m–1km	4
Less than 100 metres	2	More than 1 km	5
100m–less than 500m	3	Don't know	9

## 3. HOW LONG DOES IT TAKE TO GET THERE, GET WATER AND COME BACK ?

No. of minutes	⌚	.....
Water on premises	888	
Don't know	999	

## 4. WHAT KIND OF TOILET FACILITY DOES YOUR HOUSEHOLD USE ?

Flush to sewage system	1	Covered by dry latrine	4
Flush to septic tank	2	Uncovered latrine	5
Pour flush latrine	3	No facilities	9

**GO TO NEXT MODULE**

## 5. HOW FAR IS THE FACILITY FROM YOUR DWELLING ?

In dwelling	1	50m or more away	3
Less than 50m away	2	Don't know	9

**GO TO NEXT MODULE** ⇨**SALT IODIZATION MODULE**

**INTERVIEWER:** WE WOULD LIKE TO CHECK WHETHER THE SALT USED IN YOUR HOUSEHOLD IS IODIZED. MAY WE SEE A SAMPLE OF THE SALT USED TO COOK THE MAIN MEAL EATEN BY MEMBERS OF YOUR HOUSEHOLD LAST NIGHT?

Once you have examined the salt, complete the questions below.

1. Record test outcome:	Iodized	1
	Not iodized	2
	No salt in home	3
	Not tested	9

**GO TO NEXT MODULE** ⇨

2. Record type of salt:	Salt in bag with seal	1
	Granular (loose or coarse)	2
	Salt in blocks	3
	Other _____	4
	Not seen	9

**GO TO NEXT MODULE** ⇨

**EDUCATION MODULE**

Cluster no. \_\_\_\_ Household no. \_\_\_\_

*The questions in this module should be asked for all children in the household age 5 years and over (or over school-entry age).*

Questions	Line no.: ____ Name: _____	Line no.: ____ Name: _____	Line no.: ____ Name: _____	Line no.: ____ Name: _____
1. HAS [NAME] EVER ATTENDED SCHOOL? Yes 1 No 0 ◊ GO ON TO NEXT CHILD DK 9 ◊ GO ON TO NEXT CHILD	1 0 9	1 0 9	1 0 9	1 0 9
<b>IF THERE ARE NO OTHER CHILDREN 5 YEARS AND OVER, GO TO NEXT MODULE ◊</b>				
2. IS HE/SHE CURRENTLY AT SCHOOL THIS YEAR? Yes 1 No 0 ◊ GO TO QUESTION 4 DK 9 ◊ GO TO QUESTION 4	1 0 9	1 0 9	1 0 9	1 0 9
3. WHICH GRADE AND LEVEL IS HE/SHE CURRENTLY ATTENDING? Level: Primary 1 Secondary 2	Grade _____ ----- Level _____			
4. WAS [NAME] ATTENDING SCHOOL LAST YEAR? Yes 1 No 0 ◊ GO ON TO NEXT CHILD DK 9 ◊ GO ON TO NEXT CHILD	1 0 9	1 0 9	1 0 9	1 0 9
<b>IF NO OTHER CHILD IS 5 YEARS OR OVER, GO TO NEXT MODULE ◊</b>				
5. WHICH GRADE AND LEVEL DID [NAME] ATTEND LAST YEAR? Level: Primary 1 Secondary 2	Grade _____ ----- Level _____			

**GO TO NEXT MODULE ◊**

**MODULES FOR MOTHERS**

The following modules are directed to all mothers of under-fives in the household (page Q1, question 5 = 3). Fill in the mothers' identifying numbers and names in the spaces provided.

**TETANUS TOXOID (TT) MODULE**

Cluster no. \_\_\_\_ Household no. \_\_\_\_

Questions	Mother line no.: _____ Name: _____	Mother line no.: _____ Name: _____	Mother line no.: _____ Name: _____
1. DO YOU HAVE A CARD OR OTHER DOCUMENT WITH YOUR OWN IMMUNIZATIONS LISTED? Yes (seen) 1 Yes (not seen) 2 No 0 Don't know 9			
2. WHEN YOU WERE PREGNANT WITH YOUR LAST CHILD, DID YOU RECEIVE ANY INJECTION (E.G., TO PREVENT HIM/HER FROM GETTING CONVULSIONS AFTER BIRTH, AN ANTI-TETANUS SHOT, AN INJECTION AT THE TOP OF THE SHOULDER)? Yes 1 No 0 DK 9			
3. IF YES, HOW MANY DOSES OF TT DID YOU RECEIVE DURING YOUR LAST PREGNANCY? <b>No. of doses:</b>			
<i>If the mother reports two TT injections during the last pregnancy, STOP HERE. If she has received fewer than two TT injections during her pregnancy, continue to question. Use the following questions:</i>			
4. DID YOU RECEIVE ANY TT INJECTION (AT THE TOP OF THE SHOULDER) AT ANY TIME BEFORE YOUR LAST PREGNANCY, EITHER DURING A PREVIOUS PREGNANCY OR BETWEEN PREGNANCIES? Yes 1 No 0 DK 9			
5. IF YES, HOW MANY DOSES DID YOU RECEIVE? <b>Number of doses:</b>			
6. WHEN WAS THE LAST DOSE RECEIVED? (Record month and year OR number of years ago.)	mm/yy: or years ago:	mm/yy: or years ago:	mm/yy: or years ago:
<i>Add up responses to Q.3 and Q.5 and enter in box(es) below:</i>			
7. Total doses in lifetime:			

**CARE OF ACUTE RESPIRATORY  
ILLNESS MODULE (optional)**

Cluster no. \_\_\_\_ Household no. \_\_\_\_

Questions	Line no.: ____ Name: _____	Line no.: ____ Name: _____	Line no.: ____ Name: _____
<p>1. COUGH AND COLD ARE COMMON ILLNESSES. WHEN YOUR CHILD IS ILL WITH A COUGH AND/OR COLD, WHAT SIGNS OR SYMPTOMS WOULD LEAD YOU TO TAKE HIM/HER TO A <i>[list appropriate health providers—e.g., clinic, community health worker, doctor]</i> OR OTHER HEALTH PROVIDER? <b>Do not prompt.</b> Circle the number for <b>each</b> answer mentioned. More than one answer can be circled.</p> <p>When he/she:</p> <p>1A. has a blocked nose            1 ..... 1A</p> <p>1B. has trouble sleeping/eating    2 ..... 1B</p> <p>1C. has a fever                        3 ..... 1C</p> <p>1D. is breathing fast                4 ..... 1D</p> <p>1E. has difficulty breathing        5 ..... 1E</p> <p>1F. is ill for a long time            6 ..... 1F</p> <p>1G. other _____                7 ..... 1G</p> <p>1H. don't know                        9 ..... 1H</p>			

## MODULES FOR CHILDREN UNDER FIVE YEARS OF AGE

The following modules are directed to the mothers or caretakers of all children under age 5 in the household.

*A separate form should be filled in for each child under 5 years listed in the Household Module on page Q1. Fill in the name and line number of each child along with the cluster and household numbers in the space at the top of each module. Go through each question module with the mother. Circle the number corresponding to the mother's response where indicated. Make sure all identifying information is filled in correctly, until all children under age 5 have been covered.*

### DIARRHOEA MODULE

Cluster no. \_\_\_\_ Household no. \_\_\_\_ Child no. \_\_\_\_

Questions	Response
<p>1. HAS [NAME] HAD DIARRHOEA IN THE LAST 2 WEEKS? (<i>Diarrhoea is determined as perceived by mother, or as three or more loose or watery stools/day or blood in stool.</i>)</p> <p>Yes 1 No 0 <math>\diamond</math> <b>GO TO NEXT MODULE</b> DK 9 <math>\diamond</math> <b>GO TO NEXT MODULE</b></p>	
<p>2. DURING THIS LAST EPISODE OF DIARRHOEA, DID [NAME] DRINK ANY OF THE FOLLOWING? (<i>Prompt and circle code for all items mentioned.</i>) <b>1=Yes 2=No 9=DK</b></p> <p style="text-align: right;">Y N DK</p> <p>2A. breast milk? ..... 2A. 1 0 9</p> <p>2B. cereal-based gruel or gruel made from roots or soup? ..... 2B. 1 0 9</p> <p>2C. other locally-defined acceptable home fluids (e.g., SSS, yogurt drink)? ..... 2C. 1 0 9</p> <p>2D. ORS packet solution? ..... 2D. 1 0 9</p> <p>2E. other milk or infant formula? ..... 2E. 1 0 9</p> <p>2F. water with feeding during some part of the day? ..... 2F. 1 0 9</p> <p>2G. water alone? ..... 2G. 1 0 9</p> <p>2H. defined "unacceptable" fluids ..... 2H. 1 0 9</p> <p>2I. nothing <math>\diamond</math> <b>GO TO QUESTION 4</b> ..... 2I. 1 0 9</p>	
<p>3. DURING [NAME]'S DIARRHOEA, DID HE/SHE DRINK MUCH LESS, ABOUT THE SAME, OR MORE THAN USUAL?</p> <p>Much less or none 1 About the same (or somewhat less) 2</p> <p>More 3 Don't know 9</p>	
<p>4. DURING [NAME]'S DIARRHOEA, DID HE/SHE EAT LESS, ABOUT THE SAME, OR MORE FOOD THAN USUAL? (<i>If less, probe: MUCH LESS OR A LITTLE LESS THAN USUAL?</i>)</p> <p>None 1 Much less 2 Somewhat less 3</p> <p>About the same 4 More 5 Don't know 9</p>	



**VITAMIN A MODULE**

Cluster no. \_\_\_\_ Household no. \_\_\_\_ Child no. \_\_\_\_

*For the questions in modules A, B and C, below, fill in the appropriate number in the response column at right.*

<b>MODULE A (for countries with supplementation programme)</b>	<b>Response</b>
<p>A1. HAS [NAME] EVER RECEIVED A VITAMIN A CAPSULE (SUPPLEMENT) LIKE THIS ONE? <i>(Show capsule or dispenser.)</i></p> <p>Yes 1 No 0 ◊ <b>GO TO NEXT MODULE</b> DK 9 ◊ <b>GO TO NEXT MODULE</b></p>	
<p>A2. HOW MANY MONTHS AGO DID [NAME] TAKE THE LAST CAPSULE? <b>(DK = 99)</b></p>	
<b>MODULE B (for countries with food fortification programme)</b>	
<p>B1. WE WOULD LIKE TO KNOW IF SOME FOOD PRODUCTS ARE USED IN YOUR HOUSEHOLD. DO YOU HAVE [FORTIFIED FOOD PRODUCT] IN THE HOUSE? WOULD YOU SHOW US?</p> <p>Yes (seen) 1 Yes (not seen) 2 No 0 ◊ <b>GO TO NEXT MODULE</b></p>	
<p>B2. SINCE LAST [DAY OF THE WEEK], DID [NAME] EAT [NAME OF FOOD FORTIFIED BY PROGRAMME]? <i>(Show product package and <b>prompt</b>: USED IN COOKING, STIRRED IN DRINKS, ETC?)</i></p> <p>Yes 1 No 0 Don't know 9</p>	
<b>MODULE C (for countries with dietary education programme)</b>	
<p>C1. HAVE YOU HEARD ANY [COUNTRY-SPECIFIC] MESSAGES WHICH PROMOTE CERTAIN FOODS THAT ARE IMPORTANT FOR SIGHT AND HELP PREVENT BLINDNESS?</p> <p>Yes 1 No 0 ◊ <b>GO TO NEXT MODULE</b> DK 9 ◊ <b>GO TO NEXT MODULE</b></p>	
<p>C2. CAN YOU TELL ME WHAT SOME OF THESE FOODS ARE? <i>Circle code if mentioned. Do not prompt. (List Vitamin A-rich foods that are country/region/season-specific. The foods need not be limited to two types, as in this example.):</i></p> <p><u>Mentioned:</u></p> <p>C2A. Food 1 (country-specific) Yes 1 No 0</p> <p>C2B. Food 2 (country-specific) Yes 1 No 0</p> <p>C2C. Other responses _____ Yes 1 No 0</p>	
<p>C3. SINCE LAST [DAY OF THE WEEK], DID [NAME] EAT ANY OF THE FOLLOWING FOODS? <i>(List country/region/season-specific target Vitamin A food source.)</i></p> <p>C3A. Food 1 (country-specific) Yes 1 No 0 DK 9</p> <p>C3B. Food 2 (country-specific) Yes 1 No 0 DK 9</p>	

**BREASTFEEDING MODULE (optional)**

Cluster no. \_\_\_\_ Household no. \_\_\_\_ Child no. \_\_\_\_

Questions	Response
1. HAS [NAME] EVER BEEN BREASTFED? Yes 1 No 0 $\diamond$ GO TO NEXT MODULE OR QUESTION 4, IF INCLUDED DK 9 $\diamond$ GO TO NEXT MODULE OR QUESTION 4, IF INCLUDED	
2. IS HE/SHE STILL BEING BREASTFED? Yes 1 No 0 $\diamond$ GO TO NEXT MODULE OR QUESTION 4 AND/OR 5, IF INCLUDED DK 9 $\diamond$ GO TO NEXT MODULE OR QUESTION 4, IF INCLUDED	
3. SINCE THIS TIME YESTERDAY, DID HE/SHE RECEIVE ANY OF THE FOLLOWING? <i>Prompt and circle code for all items mentioned. 1=Yes 2=No 9=DK</i> 3A. vitamin, mineral supplements or medicine ..... 3B. plain water ..... 3C. sweetened, flavoured water or fruit juice or tea or infusion ..... 3D. oral rehydration solution (ORS) ..... 3E. tinned, powdered or fresh milk or infant formula ..... 3F. any other liquids (specify: _____) ..... 3G. solid or semi-solid (mushy) food ..... 3H. received ONLY breast milk .....	Y N DK 3A. 1 0 9 3B. 1 0 9 3C. 1 0 9 3D. 1 0 9 3E. 1 0 9 3F. 1 0 9 3G. 1 0 9 3H. 1 0 9
4. <b>Optional question:</b> SINCE THIS TIME YESTERDAY, HAS [NAME] BEEN GIVEN ANYTHING TO DRINK FROM A BOTTLE WITH A NIPPLE OR TEAT? Yes 1      No 0      DK 9	
5. <b>Optional question</b> (for countries where breastfeeding durations are very short—i.e., less than 6 months): IF [NAME] IS NO LONGER BREASTFED, AT WHAT AGE WAS BREASTFEEDING STOPPED? (Record age in months. If the mother does not know, record 99.)	

**IMMUNIZATION MODULE**

Cluster no. \_\_\_\_ Household no. \_\_\_\_ Child no. \_\_\_\_

*If an immunization card is available, copy the dates for each type of immunization below. If no date for vaccination is recorded on the card, or if no card is available, use probing questions to find out if the child received that vaccination, and if so, how many doses. Record the mother's response for each vaccine dose in the space provided.*

Questions	Y=1 N=0 DK=9	Date of immunization (day) (month) (year)			
1. IS THERE A VACCINATION RECORD CARD FOR [NAME]? Yes 1 No 0 DK 9					
2. BCG Yes 1 No 0 DK 9					
3A. DPT1 Yes 1 No 0 DK 9					
3B. DPT2 Yes 1 No 0 DK 9					
3C. DPT3 Yes 1 No 0 DK 9					
4A. OPV1 Yes 1 No 0 DK 9					
4B. OPV2 Yes 1 No 0 DK 9					
4C. OPV3 Yes 1 No 0 DK 9					
5. Measles Yes 1 No 0 DK 9					
6. BCG scar? (Check for scar, and see optional placement in the Anthropometry Module.) Yes 1 No 0 Not examined 9					
<i>Probing questions to use when no vaccination card is available:</i>					
2. HAS [NAME] EVER BEEN GIVEN A BCG VACCINATION AGAINST TUBERCULOSIS—THAT IS, AN INJECTION IN THE LEFT SHOULDER THAT CAUSED A SCAR?					
3. HAS [NAME] EVER BEEN GIVEN “VACCINATION INJECTIONS”—THAT IS, AN INJECTION IN THE THIGH OR BUTTOCKS—TO PREVENT HIM/HER FROM GETTING TETANUS, WHOOPING COUGH, DIPHTHERIA? HOW MANY TIMES?					
4. HAS [NAME] EVER BEEN GIVEN ANY “VACCINATION DROPS” TO PROTECT HIM/HER FROM GETTING DISEASES—THAT IS, POLIO? HOW MANY TIMES HAS HE/SHE BEEN GIVEN THESE DROPS?					
5. HAS [NAME] EVER BEEN GIVEN “VACCINATION INJECTIONS”—THAT IS, A SHOT IN THE ARM, AT THE OF 9 MONTHS OR OLDER)—TO PREVENT HIM/HER FROM GETTING MEASLES?					

**ALTERNATIVE IMMUNIZATION MODULE**

Record all doses of each vaccine that a child received, indicating the source of information by filling in the number of doses in the appropriate box at right. If there is no card, or if no dose is recorded on the card, ask the mother how many doses of each vaccine the child received.

Questions	Source:	
	Card	Mother's report
1. HAS [NAME] EVER BEEN GIVEN A BCG VACCINATION AGAINST TUBERCULOSIS—THAT IS, AN INJECTION IN THE LEFT SHOULDER THAT CAUSED A SCAR? Yes 1      No 0      DK 9		
2. HAS [NAME] EVER BEEN GIVEN "VACCINATION INJECTIONS" (THAT IS, AN INJECTION IN THE THIGH OR BUTTOCKS) TO PREVENT HIM/HER FROM GETTING TETANUS, WHOOPING COUGH, DIPHTHERIA? HOW MANY TIMES?	Number:	Number:
3. HAS [NAME] EVER BEEN GIVEN ANY "VACCINATION DROPS" TO PROTECT HIM/HER FROM GETTING DISEASES (I.E., POLIO)? HOW MANY TIMES HAS HE/SHE BEEN GIVEN THESE DROPS?	Number:	Number:
4. HAS [NAME] EVER BEEN GIVEN "VACCINATION INJECTIONS"(THAT IS, A SHOT IN THE ARM, AT THE AGE OF 9 MONTHS OR MORE) TO PREVENT HIM/HER FROM GETTING MEASLES?	Number:	Number:
5. BCG scar? (Check for scar.) Yes 1      No 0      Not examined 9		

**ANTHROPOMETRY MODULE**

Cluster no. \_\_\_\_ Household no. \_\_\_\_ Child no. \_\_\_\_

**INTERVIEWER:** After questionnaires for all children are complete, the measurer weighs (and **optionally** measures) each child. Record weight (and height, if measured) below, taking care to record the measurement on the correct questionnaire for that child (**check child's name**).

Measurements	Result
<p><b>Optional placement:</b> (See Immunization Module, question 6 or Alternative Immunization Module, question 5.)</p> <p>1. BCG scar? (Check for scar.)                      Yes 1      No 0      Not examined 9</p>	
2. Weight (kg) .....	□□.□
3. ( <b>Optional</b> ) Height/length (cm) .....	□□□.□
3A. Measurement made:    lying down 1            standing 2	
4. Measurer code	
5. Result: Measured 1      Not present 2      Refused 3      Other 9	

**GO ON TO NEXT CHILD** ◇

When all children in the household have been weighed, check that all modules have been filled in and that the identifying number for the household is at the top of each page. Clip all the pages together as instructed. Go back to page 1 and record that the interview is complete.

**Thank the mother for her cooperation.**

## MORTALITY MODULE

To include mortality measurement in multiple-indicator surveys, use the alternative modules below.

### HOUSEHOLD MODULE

**INTERVIEWER:** *Begin by introducing yourself—for example, WE ARE FROM \_\_\_\_\_ AND WOULD LIKE SOME INFORMATION THAT WILL HELP US IMPROVE THE HEALTH AND WELL-BEING OF CHILDREN. THE QUESTIONS WILL TAKE ONLY A FEW MINUTES.*

#### Household Information Panel

Cluster number:	Household number:	Date of interview (day/month/year): __ / __ / __
Interviewer no.:	Name of head of household:	Call-back necessary? Yes / No Time: ____ A.M./P.M.
No. persons in HH usually resident:	Material of dwelling floor: 1 wood/tile 2 planks/concrete 3 dirt/straw 4 other	Number of rooms in dwelling:
Data entry clerk no.:	All forms completed? 1 Yes 0 No	Region 1 2 3 4
	If not, why not? 1 Refusal 2 Not at home 3 HH not found/destroyed 4 other	Urban 1 Rural 2

**INTERVIEWER:** CAN YOU TELL ME THE NAMES AND AGES OF ALL WOMEN OVER AGE 15 WHO SLEPT HERE LAST NIGHT? *Fill in the following information about each woman<sup>9</sup> in the household, listing the first woman's name in line 10. First ask for names, and then go back and ask for each woman in turn, naming her: HOW OLD ARE YOU? HAVE YOU EVER HAD A CHILD? If she answers yes, ask about own-born children living with her, own-born children living elsewhere and own-born children who are now dead. (For this part of the questionnaire, you need to make it clear that you are **not** interested in the children of **other** women (i.e., children adopted or fostered into the household). If the woman is not present, ask the head of household or other women to supply the information for her. If you cannot obtain the answer to a question, **do not leave a blank space: put a 99 in the space provided for the answer.***

#### Mother Listing Form

Line no.	1. Mother's name	2. Age in years		Children ever-born to mother:						
				Living at home		Living elsewhere		Died		Total children
				Male	Female	Male	Female	Male	Female	
10										
20										
30										
40										
50										

<sup>9</sup>Eligibility criteria will vary from country to country. Where almost all childbearing occurs within marriage, only ever-married women (i.e., currently married, separated, widowed or divorced) women need to be interviewed. In such cases, it is preferable to ask for "years since first marriage" instead of age.

Fill in the Pregnancy History Form below for every mother listed in the Mother Listing Form on page Q12. If the woman is not at home, you may ask for the information about her from another woman living in the household—for example, her mother, adult daughter, co-wife, etc. Complete the Child Listing Form for every mother or other woman who is a primary caretaker in the household, even if she has not given birth herself.

Cluster no. \_\_\_\_ Household no. \_\_\_\_

**INTERVIEWER:** I WOULD LIKE TO ASK ALL MOTHERS OR OTHERS WHO CARE FOR CHILDREN SOME QUESTIONS ABOUT THE HEALTH AND WELL-BEING OF THE CHILDREN IN THIS HOUSEHOLD. For each woman with at least one birth (alive or dead) listed on the previous table, fill in the following form. Fill in one form for each woman. Ask each woman in turn about her last three pregnancies: FIRST, I WOULD LIKE TO RECORD SOME INFORMATION ABOUT YOUR MOST RECENT PREGNANCIES. WAS THE OUTCOME OF YOUR LAST PREGNANCY A LIVE BIRTH, A STILLBIRTH OR A MISCARRIAGE? Make sure the woman understands that a live birth is any child who breathed or cried after birth, even if he/she lived only a short time. **Fill in the table below, once for each woman, going back as far as three pregnancies.**

**Pregnancy History Form**

Woman's line no.: ____ Name: _____ _____	Pregnancy outcome <sup>a</sup>	Child's name	For live births only:						Age at death (months)		
			Child's sex: 1=Male 2=Female		Date of birth: (dd / mm / yy)			Still alive? 1=Y 0=N			
1. Last			1	2						1	0
2. Next-last			1	2						1	0
3. Second-last			1	2						1	0

<sup>a</sup>1 = live birth, 2 = stillbirth, 3 = miscarriage/abortion.

**INTERVIEWER:** Now fill in the Child Listing Form below about **only** the children under age 15 for whom this woman is the primary caretaker (that is, her **living** children and any other children in the household that she cares for, even if they are not the biological children of any woman in the household—i.e., fostered children, stepchildren, younger siblings of the woman). Stop listing when you reach a child over age 15. Then ask: ARE THERE ANY OTHER CHILDREN YOU CARE FOR WHO LIVE HERE, EVEN IF THEY ARE NOT AT HOME NOW? These may include children in school or at work. **If yes, list these children. Add continuation sheet if not enough room on this page. If continuation sheet used, tick here.**

**Child Listing Form**

Woman's name: \_\_\_\_\_ Line no.: \_\_\_\_\_

Child line no.	1. Child's name:	2. Child's sex: 1=Male 2=Female	3. Day, month, year of birth? (d d / m m / y y)				4. Age in years	5. Age 5 or over? Y=1, N=0
		1 2						1 0
		1 2						1 0
		1 2						1 0
		1 2						1 0

Continue to complete a form for every woman in the household who is a primary caretaker, using one page for each woman. (If the woman has never given birth, omit the Pregnancy History Form by drawing a line through it.) Copy line numbers for all children over age 5 to the next page. Copy line numbers for all under-fives to the separate Modules for Children Under Five Years of Age, one for each under-five.

## CHAPTER IV

# CHOOSING THE SAMPLE

*This chapter is written for survey coordinators and technical resource persons. It will enable you to:*

- ✓ Understand the basic concepts of sampling.
- ✓ Calculate the required sample size for national and subnational estimates.
- ✓ Determine the number of clusters to be used.
- ✓ Choose a sampling scheme.

### UNDERSTANDING THE BASIC CONCEPTS OF SAMPLING

In the context of multiple-indicator surveys, sampling is a process for selecting respondents from a population. In our case, the respondents will usually be the mothers, or caretakers, of children in each household visited,<sup>1</sup> who will answer all of the questions in the Child Health modules. The Water and Sanitation and Salt Iodization modules refer to the whole household and may be answered by any adult. Questions in these modules are asked even where there are no children under the age of 15 years.

In principle, our survey could cover *all* households in the population. If all mothers being interviewed could provide perfect answers, we could measure all indicators with complete accuracy. However, interviewing all mothers would be time-consuming, expensive and wasteful. It is therefore necessary to interview a *sample* of these women to obtain *estimates* of the actual indicators.

The difference between the estimate and the actual indicator is called *sampling error*. Sampling errors are caused by the fact that a sample—and not the entire population—is surveyed.

Sampling error can be minimized by taking certain precautions:

- Choose your sample of respondents in an unbiased way.
- Select a large enough sample for your estimates to be precise.

Choose your sample from all the households. Avoid choosing samples which might result in *biased* estimates.

---

<sup>1</sup>A *household* is a group of persons who live and eat together.



 **EXAMPLE:**

Do not choose samples exclusively from particular groups, such as children coming to clinics.

Do not ask mothers to bring their children to a central point in the community, because some of them will not come; you will not be able to find out how many failed to appear and how different they may be from those who came.

Do not use samples chosen at will by the interviewer, field supervisor or field director.

Do not restrict your sample to families living in easily accessible households, such as those close to a main road or near a village center; families living in less accessible areas may be poorer and less healthy.

Do not omit households where no one is at home the first time you call. Find out if the household is inhabited, and revisit at a later time.

If a sample does not accurately represent the whole population of interest, the estimate will be *biased*. That is, the estimate will be shifted in one direction or another from the true value.

Probability sampling ensures that all individuals in the target population (for example, all mothers of young children) have a known chance of being interviewed.

---

**To avoid bias you should use *probability sampling* to select your sample of respondents. Bias depends on the selection procedure, not on sample size. Larger samples tend to be more precise but are not necessarily less biased.**

---

In order to take a probability sample, you must have a *sampling frame*. This is a list of all the units (states or provinces, districts, communities, households, individuals, etc.) in the population from which you will choose your sample. Virtually no country has updated listings of all individuals and only very few have household listings. Your sampling frame, therefore, will typically include listings of larger population units such as states or provinces, districts, towns, villages or census enumeration areas. You must also have an estimate of the population or number of households for each of these units as well as ensuring that *all* the units are included in the sampling frame.

Since in multiple indicator surveys we are usually interested in obtaining national-level indicators, it is necessary to employ *multi-stage sampling*. For example, you might start by selecting a number of provinces (first stage), then a number of districts in each of the chosen provinces (second stage), a number of communities in each selected district (third stage) and, finally, a number of households in each chosen community (fourth stage). Each of these selections may use one of the different types of sampling listed below. Further on in this chapter, we will provide some actual examples of how different types of sampling were combined in multiple stages.

### Simple Random Sampling

Simple random sampling is the simplest form of probability sampling. Random numbers are chosen using a calculator, a computer program or a random number table. Alternatively, the names or identification numbers of all communities, households or individuals could be written on pieces of paper and the desired sample could be selected by picking the required number of papers. Each community, person or household corresponding to the numbers chosen is then included in the sample. In simple random sampling, the selection of one individual is *independent* of the selection of another individual.

### Systematic Sampling

Systematic sampling is a modification of simple random sampling. It consists of picking every  $n$ th (5th, or 10th, or 50th and so on) household from a complete list of households. When employing systematic sampling, you must ensure that the list you use is not ordered in any regular way which would bias your sample.

 **EXAMPLE:**

If the list is ordered village head, subhead, other households, then another village head, subhead, other households, and so on, systematic sampling may bias your results. For instance, if you always start at the fifth household in each village you will never include any village heads or subheads.

### Cluster Sampling

As already noted, in countries where surveys are necessary to measure goal indicators, lists of households or individuals are not usually available. It is usually too expensive and time-consuming to construct such a list and then to locate the households. The logical choice, then, is to use cluster sampling—that is, to select groups of households which are geographically close to one another.

Cluster sampling usually involves at least two stages of selection.

**EXAMPLE:**

A number of communities<sup>2</sup> may be initially selected from a sampling frame (list of communities), and then a cluster is randomly chosen within each of these communities. All households in those clusters are then interviewed.

Households are the *basic sampling unit* (that is, the smallest unit to be sampled). One might also use children under age five as the basic sampling unit, but households are much easier to identify and to list than children under five years of age. The final choice of households is made in the community. Different ways to choose households within a community are discussed in chapter 6.

Cluster sampling has many advantages. It reduces time and travel costs and simplifies field work. It also eases field supervision and survey administration. This is important since better supervision of interviewers will result in better data quality. However, cluster sampling implies that each respondent is not chosen independently of the other respondents—after all, they are neighbors—and this may increase the sampling error.

---

---

**Cluster sampling is always used in multiple-indicator surveys.**

---

---

**Sampling with Probability Proportional to Size**

Communities may vary considerably in population. If simple random or systematic sampling is used, both large and small communities will have the same probability of being included, which is incorrect. One way of compensating for these differences is to choose clusters from the sampling frame with probability proportional to size (PPS).

One advantage of PPS is that, if properly used, each household in the sample will have an equal chance of being selected. The sample is then said to be *self-weighting*, which will simplify the analysis.

---

---

**Communities with larger populations should have a proportionately greater chance of containing a selected cluster than smaller communities.**

---

---

The procedure for PPS sampling is described in the following section. You may use PPS selection at most stages of your sampling scheme. For example, you may use PPS sampling to select states or provinces in a country,

---

<sup>2</sup>A *community* is defined as a grouping within a population, such as a town, village or census enumeration area. In this manual we use the term *cluster* to mean a randomly selected sample from the community. This is the EPI usage of the term; in statistical literature, the cluster *is* the community, and other publications might refer to it in that way.

municipalities within a state or communities or villages within a district. The example below deals with selecting communities from a district but the procedure is exactly the same whatever the sampling units you choose.

### ***Procedure for Sampling with Probability Proportional to Size***

To select the communities to be used in the survey you will need:

- a list of all communities in the district to be surveyed (in this example, every household in the district belongs to one—and only one—community)
- an estimate of population size (or the number of households) in each community<sup>3</sup>
- the number of clusters you want to have in the district (discussed later)

To sample the communities with PPS, complete the following seven steps:

**1. Make a table with three columns, as shown in the illustrative table on the following page.**

Column (1): Assign a number to each community. List the communities.

Column (2): List the population of each community.

Column (3): List the cumulative population of each community—that is, the sum of the population of that community plus the populations of all the communities above it in the table.

☞ When listing the communities, try to use a logical order, such as a geographical pattern. See the discussion about *implicit stratification*, below.

**2. Calculate the *sampling interval* using the following formula:**

$$\textit{Sampling interval} = \textit{cumulated total population} \div \textit{number of clusters required}$$

Using the example above, and supposing we need to select three clusters, this would give:

$$\textit{Sampling interval} = 6,700 \div 3 = 2,233$$

---

<sup>3</sup>If the average size of households does not vary greatly from one community to another, you may use the following formula to estimate the population in each community: *population size = number of households × average size of a household*.

**ILLUSTRATIVE TABLE:  
Cumulating Community Populations**

Village	Population size	Cumulative
1	1,000	1,000
2	400	1,400
3	200	1,600
4	300	1,900
5	1,200	3,100
6	1,000	4,100
7	1,600	5,700
8	200	5,900
9	350	6,250
10	450	6,700

3. **Select a random number which is equal to or less than the sampling interval.** For the above example, we need to choose a random number between 1 and 2,233. You can select this random number by using a calculator, a computer, a table of random numbers from a statistics textbook or even by asking someone (who is not familiar with the survey) to choose a number within this range. Let us suppose that the chosen number is 1,814.

4. **Look back at the table.** Locate the first cluster by finding the community whose cumulative population exceeds this random number. In the example, the first cluster would be located in village 4, where the cumulative population is 1,900.

5. **Add the sampling interval to the random number.** In the example,  $2,233 + 1,814 = 4,047$ .

6. **Choose the community whose cumulative population just exceeds this number.** The second cluster will be located in this community. In the example, the second cluster will be located in village 6.

7. **Identify the location of each subsequent cluster** by adding the sampling interval to the number which located the previous cluster. Stop when you have located as many clusters as you need.

If a community has a population which is larger than the sampling interval, it may be

selected twice. Two independent clusters must then be selected from that same community. This is perfectly valid.

***Once you have done your PPS sampling and ensured that there are no errors, resist the temptation to do it again if your first sample does not please you.***

### **Stratified Sampling (Stratification)**

You may want to measure indicators for subgroups within your national population. Since some of these subgroups may have fewer individuals than others, you will want to ensure that all of them are adequately represented in your sample. Using PPS may result in small samples for groups representing a small proportion of the overall population. Stratification solves this problem. You will create one stratum for each subgroup and sample each stratum separately.

Stratification is commonly used to produce subnational data—for example, at the state or province level. If you do not stratify, your large states or provinces will have samples many times larger than the small states or provinces. To achieve equal precision for all states (see discussion below) you may select the same number of households in each state, and later weight the results according to the actual proportion of the national population in each state (see chapter 7).

In addition to providing subnational indicators, stratification can be used to generate data for subgroups living under specific conditions.

#### **EXAMPLE:**

Urban slum dwellers may constitute only a small fraction of the population in some countries but may present special needs. Therefore, at the beginning of the sampling procedure one might list urban slum areas separately from other areas and ensure that a sufficient number of slum dwellings are included in the sample by giving them a higher probability of selection. *This sample would then have two strata: slum dwellers and all others.*

When stratification is used, the sample size in each stratum is not usually proportional to its actual population. This imbalance has to be corrected in the data analysis. For example, if slum dwellers constitute 5 per cent of a country's population but your stratification procedure results in a sample with 10 per cent belonging to that category, the sample data will have to be *weighted* to produce national results.<sup>4</sup>

---

<sup>4</sup>In this example, each slum-dwelling child would be given a weight of 0.5 (since 5 per cent is one half of 10 per cent); also, non-slum dwellers would be given a weight of 0.95/0.90 because they account for 95 per cent of the population but only for 90 per cent of the sample. The general formula for the weight is equal to (proportion in the population)/(proportion in the sample). See chapter 6.

### ***Implicit Stratification***

Implicit stratification is a special form of stratification. It consists of listing the units to be sampled (for example, clusters) according to a logical order and then sampling systematically or with PPS from that listing.

**EXAMPLE:**

You could follow a geographical pattern from north to south or west to east, or you could start from the district capital and move away from it.

This ensures that the units will be more or less evenly spread within the listing and avoids the possibility that, due to chance, one type of community ends up being under-represented. This logical order may be geographical or may follow the distribution of a relevant variable, such as proximity to health services or socioeconomic indicators.

In addition to helping produce reliable results at the subnational level, stratification is a useful tool for improving the overall precision of your national estimate.

*The above-discussed types of sampling—simple random, systematic, cluster, PPS and stratification—are by no means exclusive. As the examples below illustrate, they may be combined at different stages of sampling to produce national estimates.*

## **HOW TO CALCULATE THE SAMPLE SIZE**

Sample size estimation is a simple procedure, yet many people are put off by its apparent complexity. This section will lead you through the various steps required for this calculation. At the end, a spreadsheet program is introduced which carries out these calculations. Understanding the basic principles of sample size estimation, presented below, is essential for the proper use of the spreadsheet.

Surveying a sample of children will give you an *estimate* of the actual goal indicator, such as the proportion of households using iodized salt. Due to sampling error, this estimate is not an exact measurement. The use of probability sampling, however, allows us to calculate a range in which the actual indicator can be reasonably expected to fall. The upper and lower limits of this range are known as the *margin of error*.<sup>5</sup>

**EXAMPLE:**

The classical EPI surveys estimated vaccine coverage with a margin of error of plus or minus 0.10, or 10 percentage points. This means that, if our survey found a BCG

---

<sup>5</sup>The term “margin of error,” used in this document, may be called “level of precision” or “confidence interval” in other publications.

coverage of 75 per cent, we could be reasonably confident that the actual population value would fall between 65 and 85 per cent.


By "reasonably confident," we mean that there will be only a 5 per cent chance that the actual coverage in this population is outside the margin of error determined by the survey. In other words, we can be confident that in 95 out of 100 surveys the true rate in the population would lie within this margin.

For sample size estimation, you must obtain or estimate some basic figures. Follow the steps below to calculate the required sample for each goal indicator. We will use BCG vaccine coverage as an example. Remember that *in the calculations you must always enter percentages as decimal numbers* (for example, enter 0.6 for 60 per cent).

**1. Guess/anticipate the *proportion* you are about to measure.** Using any available information, make a rough estimate of the proportion of children who are vaccinated. This is called the *anticipated proportion*.

 **EXAMPLE:**

You may guess that 80 out of every 100 children born have been vaccinated. Therefore, your anticipated proportion will be 0.8.

 You probably do not have good estimates of current coverages. (That is exactly the reason you are doing the survey!) However, even approximate information is adequate. When in doubt, err toward 0.5 because estimates close to 0.5 require larger samples for a given margin of error than either larger or smaller coverages.

 **EXAMPLE:**

If you think coverage is between 60 and 70 per cent, choose 0.6.

**2. Decide on the acceptable *margin of error* for your estimate.** That is, decide how close to the actual rate your estimate should be. For national goal monitoring, a reasonable margin of error is plus or minus 5 percentage points,<sup>6</sup> but for subnational estimates you may be satisfied with 10 percentage points.


---

<sup>6</sup>In this text, margins of error are expressed in terms of an absolute number of percentage points. For example, a coverage estimated at 60 per cent with an error of plus or minus 5 percentage points means that the actual coverage should lie between 55 and 65 per cent.



 **EXAMPLE:**

You may want to estimate BCG coverage within plus or minus 5 percentage points. Therefore, if your surveys finds a coverage of 82 per cent, the actual population coverage is likely to be between 77 and 87 per cent.

-  The smaller the acceptable margin of error, the larger the sample size used in the survey must be.

You must ensure that your sample is large enough to achieve the absolute margin of error that you require. For national goal monitoring, margins of error of, say, 15 percentage points or greater are probably too wide to be of any use. But you must also balance the need for precision against the costs involved in doing a very extensive survey.

 **EXAMPLE:**

If you want to estimate BCG coverage with an error of plus or minus 1 percentage point, you would require over 10,000 children aged 12–23 months. This would imply visiting 70,000 households in a typical developing country, a task which most countries would not dare to attempt.

3. **Choose an appropriate *design effect*.** When cluster sampling is used, respondents are not chosen completely independently of the other respondents in the same neighborhood.

 **EXAMPLE:**

If children within a cluster are like each other with regard to their vaccine status, we say they are "homogeneous." This means that if one given child has been vaccinated it is likely that his or her neighbors will also have been vaccinated. Homogeneity also applies to unvaccinated children: other children in the same cluster will also be likely to be unvaccinated.

If children are more homogeneous (or alike) within a cluster than in the whole population, it will result in a larger sampling error than if the same children had been obtained through a strict random sample—that is, if they had been selected completely independently of each other.

 **EXAMPLE:**

A sample of 1,000 children obtained through cluster sampling would lead to a less precise coverage estimate than another sample of 1,000 children chosen through simple random sampling.

Some indicators will be more affected by clustering than others. Vaccine coverage may be affected because of distance from health centers and by local immunization campaigns.

Homogeneity is particularly severe for indicators of incidence of infectious diseases such as measles (which is spread from child to child) as well as variables related to water and sanitation (since neighbors often share the same sources of water and sanitation facilities).

The mathematical expression of this clustering is called the *design effect*. It depends both on the degree of similarity among children within a cluster and on the size of the clusters. Other things being equal, large clusters will lead to a greater design effect than small ones.

Since sample size calculations are usually based on simple random samples, the size of the sample must be increased to compensate for a large design effect. Using experience from other cluster surveys, we usually allow for a design effect of 2.0 for most variables, which requires us to double the sample size in comparison with a simple random sample. For water and sanitation variables, we have used a design effect of 10.

☞ It is very hard to predict what your design effect will be before carrying out the study. Unless you have good reason to the contrary, use the figures suggested above. After your survey is finished, you can calculate the design effect (see chapter 7) and use it to calculate your actual margin of error.

#### 4. Calculate the sample size using the formula:

$$\text{sample size} = \frac{4 \times \text{proportion} \times (1 - \text{proportion}) \times \text{design effect}}{\text{margin of error} \times \text{margin error}}$$

#### 📎 EXAMPLE:

Using this formula, and the values given in the examples above for BCG coverage, where:

$$\text{proportion} = 0.8 \text{ (that is, 80 per cent)}$$

$$\text{design effect} = 2$$

$$\text{margin of error} = 0.05 \text{ (}\pm 5 \text{ percentage points)}$$

we have:

$$\text{sample size} = \frac{4 \times 0.8 \times (1 - 0.8) \times 2}{(0.05 \times 0.05)} = 512$$

So the sample size should be 512 children aged 12–23 months. That is, to obtain an estimate with a margin of error of plus or minus five percentage points (a 95 per cent chance that the true rate falls within 75 per cent and 85 per cent), you should interview mothers of 512 one-year-old children.

These calculations must be repeated for each of the indicators being measured in the survey. This can be done with the help of a spreadsheet program (see Box 4.1, beginning on page 4.14).

**5. Estimate the number of households you must visit.** The above formula and the spreadsheet provide the number of members of the target population required for measuring the indicator. Note that the target populations vary according to each indicator—for example, children aged 12–23 months for vaccine coverage, children with recent diarrhoea for oral rehydration therapy (ORT) use rate, households for water and sanitation, and so on—so that a common denominator is necessary.

- ☞ By calculating the number of households that you must visit to find the required number in the target population, the sample size requirements for all indicators may be compared. The largest required number of households should then be chosen. This guarantees that all other indicators will also be measured with equal or smaller margins of error.

To estimate the number of households required for finding a given number of children, you must know the average household size and the proportion of children in the population. Note that the larger the households and the greater the birth rate, the fewer households you will have to visit to find a sufficient number of children in any age range.

**EXAMPLE:**

If the average household size is six persons, and 3 per cent of the population is made up of children aged 12–23 months, in order to find 512 children in this age range you must visit  $512 \div (6 \times 0.03) = 2,844$  households.

- ☞ Check the last available census to find out the average household size and the proportion of children under 5 years of age in the population. If you do not have such data, typical figures are six persons per household and 3 per cent of the population being children in each one-year age group (e.g., aged 0–11 months, 12–23 months, etc.), so that 15 per cent of the population will consist of under-fives. You may also use 25 per cent as the frequency of diarrhoea in the preceding two weeks. Table 4.1 provides guidance on estimating the number of households for measuring the mid-decade goal indicators.

**6. Allow for possible losses.** In all surveys, some of the planned interviews fail to take place because people are absent from their homes and because of refusals. Strategies

**Add at least 10 per cent to your estimated sample size to allow for losses.**

for minimizing these losses are discussed in chapter 6. When calculating the sample size, however, it is wise to add to the sample the expected proportion of losses. These may vary widely from country to country, and from urban to rural areas, particularly where maternal work outside the home is common. You must estimate what this proportion is from previous surveys in the country.

**EXAMPLE:**

If your estimated sample is 2,844 households, add about 10 per cent and try to reach 3,100 households.

**Table 4.1. Typical numbers of households to be visited for finding, on average, one individual belonging to the target population for each mid-decade goal indicator**

Indicator	Target population	Per cent of whole population <sup>a</sup>	Number of households to be visited to find one individual <sup>b</sup>
TT2 coverage	Mothers of children 0–11 months	3%	5.56
Other vaccine coverages	Children 12–23 months	3%	5.56
Vitamin A coverage	Children 0–23 months	6%	2.78
Iodized salt consumption	Households	—	1.00
Per cent with low weight or height for age	Children 0–59 months	15%	1.11
Use of ORT in diarrhoea	Children 0–59 months with diarrhoea in the previous two weeks	4% (25% of 15%)	4.17
School enrollment	Children 5–9 years	15%	1.11
Water and sanitation	Whole population	100%	0.17

<sup>a</sup>Typical estimates for less developed countries.

<sup>b</sup>Assuming six individuals per household.

The discussion on sample-size estimation is followed below by three specific instances of sample sizes: for estimating trends (Box 4.2), for subnational analyses (Box 4.3) and for subgroup—for example, gender-specific—indicators (Box 4.4).

#### SUMMARY OF SAMPLE-SIZE CALCULATIONS

**Based on a number of simulations using data from different countries, measuring Mid-Decade Goals (MDGs) with a precision of 5 percentage points will require a sample of about 4,000–5,000 households.**

**The total population of your country is not important for defining the sample size.**

**Box 4.1**  
**SAMPLE SIZE ESTIMATES FOR MEASURING MID-DECADE GOALS**

**Instructions for using the Spreadsheet SAMPGOAL.WK1**

**Data Entry**

Please follow these steps to enter your data:

1. Run Lotus or Quattro-Pro (or any common spreadsheet program) and open the SAMPGOAL.WK1 file.
2. Enter data in the highlighted cells only. Note that all percentages or proportions should be entered as decimals (e.g., 15% should be entered as 0.15).
3. Provide country-specific data on the following basic assumptions:
 

**Cells C4 and D4.** These are the estimated design effects due to the use of cluster sampling. The program defaults are 2 for most variables and 10 for the water/sanitation indicators (since the latter are likely to be more affected by clustering). Unless you have local information on these, please leave them as they are.

**Cell C5.** Average number of persons per household, obtained from local census or survey.

**Cell C6.** Per cent of the population under 5 years of age, also from local census or survey.

**Cell C7.** Prevalence of diarrhoea in the preceding 15 days among children under 5 years of age. Obtain this number from previous studies in the country. If you do not have any data, use the default value of 0.25.
4. Provide the data on the estimated coverage level for each goal indicator in cells D12 through D24. You may change the indicators (columns A and B), but please make sure that they refer to the same target population (column C) or you are likely to get wrong results.
5. Enter the margin of error you are prepared to accept in cells E12 through E24. This should be expressed as percentage points.

**Results**

The spreadsheet calculates the target sample (in column F)—for example, the number of children aged 12–23 months required to measure each indicator. It also uses the information provided in the top section of the spreadsheet to calculate how many households you will have to visit (column G) to find the required number of persons in the target sample. Finally, it picks the largest of these numbers of households to be the required sample size (cell G26).

Perhaps the most useful application of this spreadsheet is to perform repeated simulations, changing the parameters in order to observe the effect on sample size. In particular, small changes in the margin of error may have important effects on the required sample.

The following table shows an example of the spreadsheet.

<b>Box 4.1 (continued)</b>						
<b>Sample Size Calculations for Measuring Mid-Decade Goals</b>						
<b>Basic Assumptions</b>		<b>Low</b>	<b>High</b>			
Design effect		2	10			
Persons per household		6				
Pct of population <5 years		0.15				
Prevalence of diarrhoea 15 days		0.25				
<b>Goal</b>	<b>Indicator</b>	<b>Target population</b>	<b>Estimated prevalence</b>	<b>Margin of error</b>	<b>Required target sample</b>	<b>Required number of households</b>
1.1	DPT3 coverage	12–23 mos	0.7	0.05	672	3,733
1.2	Measles coverage	12–23 mos	0.4	0.05	768	4,267
1.3	OPV3 coverage	12–23 mos	0.5	0.05	800	4,444
1.4	BCG coverage	12–23 mos	0.8	0.05	512	2,844
1.6	TT2 coverage (pregnancy)	0–11 mos	0.3	0.05	672	3,733
5.1	Vitamin A coverage	0–23 mos	0.3	0.05	672	1,867
6.1	Iodized salt consumption	Households	0.2	0.05	512	512
7.1	Use of ORT(1) in diarrhoea	Diarrhoea <5 yrs	0.4	0.05	768	3,413
7.2	Use of ORT(2) in diarrhoea	Diarrhoea <5 yrs	0.5	0.05	800	3,556
11.1	Percent low weight/age	All <5 yrs	0.4	0.05	768	853
12.4	School enrolment	5–9 yrs	0.7	0.05	672	747
13.1	Safe water	Population	0.6	0.05	3840	640
13.2	Sanitation	Population	0.2	0.05	2560	427
<b>Required number of households</b>						<b>4,444</b>

**Box 4.2****SAMPLE SIZE CALCULATIONS FOR MEASURING TRENDS**

The foregoing discussion refers to sample size calculations necessary for estimating a proportion within a given precision. Such calculations are adequate for many Mid-Decade and Year 2000 goals, such as reaching 80 per cent immunization and ORT use rates. Some of the goals, however, are expressed as expected reductions, such as decreasing the prevalence of malnutrition by 20 per cent between 1990 and 1995, or closing the education gap by one-third in the same five-year period. This assessment requires at least two surveys at different points in time. Since many countries do not have reliable data for 1990, they will not be able to measure these goals.

The sample size calculations for measuring such *trends* are somewhat more complex and are not be discussed in detail. There are two useful rules of thumb, however:

- If your two samples are completely unrelated (e.g., different clusters are used in the two surveys) and if your estimate has an error of plus or minus five percentage points, then the smallest trend difference you should be able to detect would be seven percentage points, or 1.4 times greater than your original margin of error.
- If you use the same clusters in both surveys, your precision improves and you should be able to measure a change of the same size (for example, five percentage points) as your margin of error.\*

Using the same clusters in repeated surveys has advantages and disadvantages. In addition to reducing sample size, it simplifies field work because the areas will already have been mapped in the first survey. On the negative side, surveys may include educational messages or raise the community's awareness of health problems and therefore might lead to changes in health behavior (known as the "Hawthorne effect"). Subsequent surveys in the same areas may, therefore, be misleading because these communities may no longer be representative of the country. When considering repeated surveys, each country should weigh these advantages and disadvantages.

**\*Technical note:** This assumes a correlation of 0.5 between the cluster-specific estimates in the first and second surveys.

**Box 4.3**  
**SAMPLE SIZES FOR SUBNATIONAL ESTIMATES**

Thus far, we have been concerned with sample sizes necessary for generating national estimates of indicators for the Mid-Decade Goals. Several of these indicators are based on specific (and often small) population subgroups such as children aged 12–23 months, or children who recently had diarrhoea. Most countries, however, will also want to use the surveys to provide subnational estimates—for example, at the level of regions, states, provinces or districts. These data may be used for detecting areas where greater efforts are required, as well as for programming and evaluation purposes.

The most limiting factors affecting national sample size are that immunization is based solely on children aged 12–23 months and that ORT use refers to children with diarrhoea in the preceding two weeks.

A possible solution is to report on indicators at subnational levels for *all under-fives* (or children aged 12–59 months for immunization)\* as well as to accept a margin of error of plus or minus 10 percentage points. This would reduce the required sample sizes at subnational levels to about 300 households per district/region. A separate version of the spreadsheet, named DISTRICT.WK1, is provided for these district-level calculations.

With 300 households, ORT use would still not be measured properly. Two options are:

1. Extend the 15-day recall period to one month, thus raising the percentage of under-fives with diarrhoea from about 25 per cent to about 40 per cent in a “typical” district. This might lead to some bias due to selective recall of more severe episodes, which tend to be more intensively treated, but it should not affect the ranking of the districts. Even with one-month recall, you would still need 500–600 households per district for a margin of error of 10 percentage points, which will require changing the first question in the diarrhoea module to refer to the last month. An extra question will be required for separating episodes starting up to 15 days before the interview (for reporting on national Mid-Decade Goals) from the remainder.
2. A second option is to ask local health planners to group districts which they believe have similar ORT use rates and to analyse the data at this multi-district level.

**By expanding the target populations for mid-decade goal indicators, you will be able to produce locally useful information for programming at low cost, and yet be able to report nationally on the narrow subgroups required.**

\*Widening the age groups may lead to some loss of validity in selected indicators due to poor recall, but this possible bias is likely to be small from the health planner's point of view. Also, this bias might be expected to operate uniformly in all districts so that ranking of areas for prioritizing interventions would not be affected.



**Box 4.4**  
**SAMPLE SIZES FOR SUBGROUP ANALYSES**

Subgroup analyses may include breakdowns of the indicators by gender, socioeconomic group and so forth. Indicators based on subgroups will be less precise than those calculated for the whole sample.

**The smaller the subgroup, the less precise the estimate.**

The examples below show how the margins of error increase for smaller subgroups. Based on an overall sample with a margin of error of plus or minus five percentage point for a given indicator,\* one would have:

- A margin of error of approximately  $\pm 6.3$  percentage points for gender-specific indicators (given 50 per cent boys and 50 per cent girls in the sample).
- A margin of error of approximately  $\pm 8.6$  percentage points for a subgroup making up 20 per cent of the sample (for example, a given socioeconomic category).

**These results show that if the overall margin of error is about five percentage points, reasonably precise results will also be obtained for gender-specific indicators as well as for other subgroups making up 20 per cent or more of the whole sample.**


\***Technical note:** Based on estimating a coverage of 50 per cent and a design effect of 2.0. Also assumes that the coverage in each subgroup will be close to 50 per cent.

### Decide on the Total Number of Clusters

A basic consideration in deciding the size of your clusters is to estimate the number of households which can be covered in one day by your team of interviewers. You do not want your field workers to spend more than a day in a given cluster, nor do you want them to finish their interviews at, say, 3 P.M. and have no further work to do for the rest of the day.

 **EXAMPLE:**

Let us assume that your interview takes 10 minutes for each child under 5 years and 3 minutes for children aged 5–9 years (education module only), plus 5 minutes for the household questions. A typical household will include one child in each of these groups (see Table 4.1) and therefore last 18 minutes. Allow 5 more minutes to move from house to house. Considering a 6-hour working day (to allow for transportation to and from the selected area and for a midday break), one interviewer would cover about 15 households a day  $[(6 \text{ hours} \times 60 \text{ minutes}) \div 23 \text{ minutes}]$ . A two-interviewer team will cover 30 households a day, which then becomes the cluster size.

 You should do a pre-test to estimate how long each interview takes. Field workers should

be well familiar with the questionnaire before the pre-test since training greatly reduces the duration of the interview. Due to distance, field work in rural areas is always much slower than in urban areas.

**Typical cluster sizes range from 20 to 40 households. Do not use a cluster greater than 100 households. It is always better to use *more small clusters* than fewer, larger clusters.**

The number of clusters is given by the formula:

$$\text{Number of clusters} = \frac{\text{Total number of households in sample}}{\text{cluster size}}$$

**EXAMPLE:**

Assuming that the interviewing team can cope with 30 households in a day, and aiming at a national sample of 4,500 households, you would need  $4,500/30 = 150$  clusters.

**Most national samples will include 100 or more clusters. For subnational estimates, try to have 15 to 20 clusters in each province, state or district. If in doubt, increase the number of clusters to be sampled.**

### Use a Flow Chart to Choose Your Sampling Scheme

When choosing a sampling scheme, it is desirable to aim at the most accurate and least biased sample. However, this choice is also affected by the availability of population data for building a sampling frame. Finally, time, feasibility and cost considerations also play a major role.

A large number of combinations of different types of sampling may be used to produce a multi-stage sample for a multiple-indicator survey. The flow chart in Figure 4.1 (see page 4.21) shows the decisions that have to be made. This example refers to a three-

**Unless the survey can be finished within a few months at reasonable cost, it cannot be used for goal monitoring and subnational planning.**

stage sample in which large areas are selected first; the second stage includes selection of small areas within the large areas; and the third stage consists of the selection of households within each of the small areas. The final alternatives, labeled A to D in the bottom right corner of the flow chart, are ranked in order of preference.

In order to design the sampling scheme, you will need the help of professionals who are familiar with the types of population data available in the country (such as census bureau officers or demographers), as well as sampling experts. If the latter are not available locally,

international help will have to be arranged.

**The ample experience with EPI surveys in the 1980s showed that a survey does not have to be perfect to produce usable data. Survey methods have to be sufficiently good for the purposes for which the results will be used.**

### *How to Select the Large Areas*

The first item on the flow chart is whether recent data are available on the population of large areas. These may be administrative divisions such as states or provinces, districts, municipalities and so on. Large areas will typically have populations ranging from several thousands to millions. To use probability sampling, it is necessary to have appropriate data on the population of these large areas—that is, a sampling frame. You have several options (in order of accuracy):

1. You can go to a recent census (say, within the last 5–10 years).
2. A recent Demographic and Health Survey (DHS) or other national household survey is another option. Although these do not usually provide population data per se, the national sampling frames have to be updated before they are carried out.
3. If no recent census or sampling frame is available but population growth since the last census has been fairly uniform throughout the country, population data from the existing census may be used.
4. If there is no recent census or other sampling frame and population growth has been variable (due to migration or disasters, for example), you should try to estimate regional population changes by consulting with local experts, including demographers, planners or workers at statistical bureaus.

**Although you should attempt to have the best possible sampling frame, it does not have to be completely accurate. Later stages of the sampling scheme may help correct some of the inaccuracies in your frame.**

Options 1 through 3 above provide data for using a large-area population to carry out the first stage of sampling procedure. Option 4, if completed successfully, will also allow it.



Once the sampling frame is available, you can achieve sampling at the first stage through stratification, through probability proportional to size (PPS) or with a combination of the two. This stage may also be subdivided into two or more substages.

 **EXAMPLE (1):**

Suppose you want separate estimates for each province in the country. You would stratify the country into its provinces and take equal-sized samples in each to obtain estimates at the provincial level. For estimating the indicators at the national level, it would be necessary to weight the provincial results according to the population in each province.

 **EXAMPLE (2):**

Suppose that you cannot afford to obtain separate estimates for each province, but that your country is divided into three broadly similar regions, the capital (with 30 per cent of the population), the coastal provinces (with 50 per cent) and the interior provinces (with 20 per cent). You might then select equal-sized samples in each region in order to obtain regional estimates. By later weighting the results according to the population, you would produce the national estimates.

In a few countries, particularly those under emergency situations, none of the options presented above for obtaining a sampling frame may be available (see flow chart). One should then work with local informants to stratify the country into broad categories of large areas. This approach does not provide a strict probabilistic sample due to the lack of population data. However, if no proper sampling frame is available, this option may at least provide approximate information on the levels of indicators.

 **EXAMPLE:**

Following the example above, the country may be divided into the capital, the coastal region and the interior region, and equal-sized samples taken in each region. Unlike the last example, it would not be possible to weight the results since you would not know the population in each stratum.

Having chosen the large areas, the next step is to select small areas.

### ***How to Select the Small Areas***

The second question in the flow chart refers to small areas. These areas include census enumeration tracts or small administrative units such as communities, villages or submunicipal units. Small areas will typically have less than 1,000 households. Data on population and boundary maps are required

if one is to use probability sampling at this stage. The chart assumes that if population data are available, boundary maps will also be present.

If a recent census was carried out, data on enumeration tracts should be available. However, these may not be readily accessible to the researchers. If available, enumeration tracts have the advantages of (1) dividing the population into mutually exclusive tracts so that each household is assigned to one and only one tract, and (2) being usually of uniform size since they are often defined as the area that one enumerator can cover.

One alternative to using enumeration tracts is to use small local administrative units, so long as these include all of the population and there is no overlap between different units.

Once information about the population of small areas has been obtained, these may be sampled with PPS. The table exemplifying PPS selection in the section above shows how villages are drawn from a district.

---

---

**If using villages as sampling units, you should ensure that every scattered rural household belongs to a given village.**

---

---

If census enumeration areas are used and they are roughly of the same size, you may use simple random sampling instead of PPS, since every household would have approximately the same probability of being selected.

If, following the flow chart, you find that data on small areas are not available, you should work with local informants to divide the large area into broad groupings of small areas, and then sample randomly from within each grouping.

 **EXAMPLE:**

Let us suppose that the large area (district) has one city as well as towns, villages and scattered rural households. Let us also suppose that there are highland and lowland populations. Strata would be created according to these characteristics, if possible, with rough population estimates. The local informants would then attempt to make a complete listing of the names of all areas (e.g., towns, villages or rural neighborhoods) in each stratum. A number of small areas—based on the crude population estimates—would then be randomly selected in every stratum.

As noted above, the use of key informants does not provide a strict probabilistic sample but may be the only alternative under some circumstances.

Having selected the small areas, it is then necessary to choose the clusters containing the households to be visited.

### ***How to Select the Clusters***

Most small areas will include several times more households than our chosen cluster size. It is then

necessary to sample households within these areas. The flow chart shows that there are four basic ways of doing this: enumeration (option A in the flow chart), segmentation (option B), or the EPI random walk method (further subdivided into options C and D). The main advantages and disadvantages of these methods are summarised in Table 4.2.

When choosing a method, you should consider its advantages and disadvantages, as well as local limitations, including availability of maps and population information, time and cost considerations, geographical characteristics and type of field personnel available.

**Table 4.2. Some advantages and disadvantages of the main methods of selecting households within a small area**

Method of selection	Advantages	Disadvantages
<b>Enumeration</b>	<ul style="list-style-type: none"> <li>• Strict probability sample.</li> <li>• Adjusts for changes in population.</li> <li>• Does not use contiguous households and thus is closer to simple random sample.</li> </ul>	<ul style="list-style-type: none"> <li>• Needs listing and mapping every household.</li> <li>• Households are not contiguous so that field work takes longer.</li> <li>• Houses must be identified precisely.</li> </ul>
<b>Segmentation</b>	<ul style="list-style-type: none"> <li>• Strict probability sample.</li> <li>• Adjusts for changes in population.</li> <li>• Saves time by using contiguous households.</li> </ul>	<ul style="list-style-type: none"> <li>• Needs extra time for sketch mapping.</li> <li>• Use of contiguous households may affect sample homogeneity.</li> <li>• Supervisors need thorough training for sketch mapping and segmenting the area.</li> </ul>
<b>Random walk</b>	<ul style="list-style-type: none"> <li>• Saves time by using contiguous households.</li> <li>• Does not require mapping.</li> <li>• Widely known and used in many countries by non-experts.</li> <li>• Particularly adequate for dispersed populations and large tracts.</li> <li>• May produce unbiased sample.</li> </ul>	<ul style="list-style-type: none"> <li>• Not a strict probability sample.</li> <li>• Does not account for changes in population.</li> <li>• Use of contiguous households may affect sample homogeneity.</li> <li>• Requires thorough training of interviewers for locating houses and close supervision.</li> </ul>

**Enumeration.** Enumeration consists of mapping and listing all households in the small area and then taking a sample, usually a systematic one. This is the method of choice for countries where listings and maps are available (for example, at health facilities). However, in the context of stand-alone multiple-indicator surveys, we do not recommend this method unless current household listings are available, because of the time required for the initial step of listing and mapping households (see Table 4.2). A "current" household listing should not be more than two years old.

**Segmentation.** The flow chart shows that the next option is segmentation. Note that, even if population data are not available but boundary maps are, segmentation may still be used.

Segmentation requires driving or walking through the whole extent of the small area and making an approximate map with the location of the households. This map does not have to be fully accurate nor is there a need to have a complete listing of all households. The map, however, should cover the whole of the small area and the segments should be mutually exclusive.

After sketch mapping, the area is divided into approximately equal-sized segments. The size of each segment is given by the formula below:

$$\text{segment size} = \frac{\text{actual number of households in area} \times \text{desired cluster size}}{\text{estimated number of households in area}}$$

 **EXAMPLE:**

Your estimate from the past census tells you that there should be 510 households in the area. Since you want clusters of, on average, 40 households, the area should be divided into 12.75 segments (or 13 in round numbers). Your sketch map shows that the number of households since the census has increased to 600. Divide the area into 13 equal-sized segments with 46 households (rounded from 46.15) each. Select one of these at random and include all households in your sample.

Since it provides automatic adjustment for inaccuracies in the sampling frame and also because it is a strict probability sample, segmentation should be used whenever it is feasible to visit the whole selected area, to draw the sketch maps and to segment the area

---



---

**A major advantage of the enumeration and segmentation methods is that they correct for inaccurate sampling frames for the small areas.**

---



---

prior to data collection. Feasibility considerations should include time, costs, geographical size and availability of experienced supervisory personnel—for example, from a department of statistics or census bureau. Chapter 6 describes in detail the steps required for segmentation.

**Random Walk Methods.** The random walk methods are based on the original EPI cluster survey method. These methods do not produce strict probability samples, but in many circumstances may provide virtually unbiased samples of households. They are also very feasible and rapid, and in some countries may constitute the only viable alternative.

Even when maps of the small areas are available, it may not be feasible to sketch-map the households. This may occur when the sampling units at this stage (for example, census tracts or villages) cover large geographical areas and/or when households are dispersed throughout the area. Under such circumstances, the mapping exercise could be very time-consuming. You would then select option C, by randomly choosing one of several starting points on the boundary map and using



that point to start the random walk. This method is further described in chapter 6.

If there are no boundary maps, a number of adaptations of the EPI random walk method may be used (option D). The original method and its limitations are discussed in Box 4.5, and chapter 6 presents a modified version of the method which overcomes some of these problems. The advantage of option C, compared with option D, is that you can use an available boundary map to allocate starting points within the whole small area. In option D, no maps are available, so you must resort to alternative ways of finding a starting point, which are described in Box 4.5 as well as in chapter 6.

## COUNTRY EXAMPLES

Experience conducting multiple-indicator surveys in three countries—Bangladesh, Brazil and Jordan—is described in the following sections.

### Bangladesh

Nationwide multiple-indicator surveys were carried out in Bangladesh in 1992, 1993 and 1994. These surveys were aimed at providing national data on mid-decade goal indicators as well as district-level data for programmatic purposes. Based on the sampling frame from the 1991 Census, the country was divided into 75 strata (or large areas): the rural areas of 64 districts and 11 urban domains including the four large metropolitan areas—two of which were further divided into slum and non-slum strata—and samples of other urban areas.

In each large area, all unions (administrative units of 3,000–4,000 households) were listed following a geographical order, and 13 were selected with PPS. For each selected union, all *mouzas* (in rural areas) or *mahallahs* (in urban areas) were listed. *Mouzas* and *mahallahs* are administrative subdivisions, each consisting of about 250–300 households; they are referred to as "small areas" throughout this Bangladesh survey description.

After the listing, one small area was selected in each union, again using PPS. Therefore, 13 small areas were chosen in each district or urban domain. In each of these 13 areas, the segmentation method was used. Most areas already had updated maps showing all households which had been drawn by health workers, so that sketch mapping was necessary in only a few instances. Each segment contained, on average, 40 households. One segment was randomly selected in each small area and all its households were included in the sample. Each district, therefore, included, on average, 520 households ( $13 \times 40$ ). This was sufficient for estimating all mid-decade goal indicators with margins of error of 10 percentage points or less.

Data from the different rural and urban samples were weighted to provide national estimates of the indicators. Since the national sample comprised 39,000 households ( $520 \times 75$ ), its margin

of error was much smaller than that of five percentage points used in the examples above. This large sample resulted from the great number of districts in the country and from the need for district-specific data. Most countries will require much smaller samples, as noted above in the section on sample size.

Another exception in the Bangladesh surveys is the use of 13 clusters per district, while above we recommended at least 20. This choice was based on practical considerations resulting from the large number of districts and urban strata, as well as on the fact that these areas were geographically small and that the characteristics of the population within each area appeared to be reasonably uniform. The national sample included 975 clusters ( $13 \times 75$ ).

**The Bangladesh example shows how it is possible to combine the process of obtaining national estimates for monitoring goals with the generation of district-level data for programming purposes.**

### **Brazil**

Multiple-indicator surveys were carried out in several of states in Northeast Brazil in the early 1990s. They were aimed at measuring child health and nutrition indicators for monitoring and programmatic purposes. The example that follows is based on the 1991 survey in the State of Pernambuco, with a population of approximately seven million.

A sample of 1,000 children under five years of age was required for measuring indicators with a margin of error of plus or minus three percentage points, using a design effect of 2.0. Based on information on household size and the proportion of the population under the age of five years, 1,920 households had to be visited. Since the questionnaire was extensive—covering children under the age of five as well as women aged 15–49 years—and the households were often dispersed, the cluster size was set at 12 households.

In the first stage of sampling, all municipalities ("large areas") were listed by starting with the capital city and following a geographical order. Using PPS, 20 numbers were selected. Since the capital city included 15 per cent of the population, it was selected three times ( $.15 \times 20$ ).

In the second sampling stage, eight census enumeration areas ("small areas") were chosen in each selected municipality through simple random sampling, totalling 160 clusters ( $20 \times 8$ ) for the state. In the capital, which had been selected three times, 24 enumeration areas ( $3 \times 8$ ) were taken. Each enumeration area includes 200–300 households, and since their size is uniform there was no need for PPS at this stage. These areas cover the whole municipality, including both urban and rural households. Four to six arbitrary starting points were plotted on the maps of each enumeration area and one of these was selected at random. Twelve households closest to this point and belonging to the enumeration area were included in the sample.

Note that this sample was self-weighting, so that there was no need to weight the estimates.

**Box 4.5**  
**THE STANDARD EPI CLUSTER SURVEY METHOD**

In the standard EPI method, 30 villages were selected with PPS within a district. A starting point was selected in a central location such as a health facility. A bottle was spun at this central point and the interviewer then counted the number of households until the edge of the village and selected one at random. This is the first household in the sample. Subsequent households are selected by going to the house whose front door is closest to that house which has just been visited, until seven children aged 12–23 months (the age range of interest for vaccine coverage) were reached. This method is known as the “random walk.”

A number of concerns were raised relative to this method, including:

1. In most countries, rural populations are not neatly organised into villages with well-defined edges or boundaries, but live in more or less scattered households.
2. The scheme may oversample households close to village centers which might not be representative of the whole village. By moving from one household to the next closest one, scattered rural households are likely to be underrepresented.
3. By stopping when reaching a fixed number of children (not households), the method oversamples areas with few children (where many households had to be visited for finding seven children) and underrepresents areas with many children (where seven children were found within a few households).
4. It is very difficult to ensure that field workers have followed the instructions regarding the “random walk.”
5. The method of selection does not guarantee that every household has a known probability of selection. Therefore, it is not a strict probability sample. This is a somewhat technical point but it has caused concern among many statisticians because, strictly speaking, you cannot calculate the margins of error of the indicators.

**These issues do not undermine the usefulness of the EPI-type survey for estimating vaccine coverage for programmatic purposes, where a wider range of errors can be tolerated. They show, however, that the method can be improved.**

### **Jordan**

A national survey for assessing immunization coverage, diarrhoea management and infant mortality was carried out in Jordan in 1991 following the EPI cluster survey methodology. A sample of 12,600 households was required for estimating cause-specific infant mortality. Note that this sample is considerably larger than that usually required for estimating vaccine or ORT coverage, but

mortality studies often require visiting many households for identifying a sufficient number of deaths.

In the first stage of sampling, the Kingdom was divided into its seven governorates. Since separate estimates were desired for each of these, the standard EPI sample of 30 clusters per governorate was adopted. Each cluster included 60 households, which ensured that the national sample would be of the required size ( $30 \times 60 \times 7 = 12,600$ ). The use of fixed samples in all governorates, independently of their population, required later weighting in the analysis.

Within each governorate, multi-stage sampling was used. First, localities were stratified into nine groups according to population size, and a number of clusters was allocated to each stratum based on its total population. The required number of localities was then selected within each stratum with PPS. Next, blocks were chosen within each selected locality and, finally, households were selected within each of these blocks.

## CHAPTER V

# PREPARING FOR DATA COLLECTION

*This section is written for survey coordinators and technical resource persons.*

*It tells you how to:*

- ✓ Make logistics arrangements.
- ✓ Prepare the questionnaire and training materials.
- ✓ Select and train the field workers.
- ✓ Choose and prepare the equipment.
- ✓ Carry out the pilot study.
- ✓ Set up computers and hire data processing staff.
- ✓ Make arrangements for returning the questionnaires to headquarters.

### **MAKING LOGISTICS ARRANGEMENTS**

In order to make logistics arrangements, you will have to (1) set up central local headquarters, (2) contact local authorities where the survey will be carried out, (3) decide on the size and composition of the field team, (4) arrange accommodations, transportation and security, and (5) arrange to obtain or prepare copies of local maps.

#### **Setting Up Central and Local Headquarters**

The study must have a central headquarters, preferably in the capital, from where the whole operation will be coordinated. It is often possible to set up this office at a government institution such as a Ministry of Health center. Usually, two to three rooms are needed for general administrative activities, in addition to meeting rooms where the training can take place. The headquarters office will hold the computing equipment and serve as the storage place for questionnaires. Telephone or radio facilities are necessary for keeping contact with the field teams.

#### **Contacting Local Authorities**

In some areas, the arrival of a team of strangers may be regarded with suspicion. In the past, the survey team's failure to contact the local authorities in advance, to let them know what they would be doing, has caused problems. In one extreme example in Latin America, the interviewers who had

not contacted the local authorities before arriving were arrested. You can avoid such problems by contacting local authorities and community leaders before beginning the study, to ask for their permission and to advise them of the team members' arrival dates. You can also ask them at this time to identify suitable local guides and, if necessary, translators. Another area in which local authorities may be helpful is in arranging accommodations and meals for the survey team.

### **Deciding on Size and Composition of Field Teams**

The number of interviewers required depends on the sample size, on the number of days to be spent interviewing and the on the number of respondents one interviewer can interview in a day. You can estimate this number from the length of a working day divided by the amount of time it takes to complete one interview (determined when you pretest the questionnaire—see below), allowing some travel time. (Travel time will be substantially longer in rural than in urban areas.)

#### **EXAMPLE:**

In chapter 4, an example was given in which each interview would last, on average, 18 minutes, plus 5 more minutes to move from house to house. Considering a 6-hour working day (to allow for transportation to and from the selected area and for a mid-day break), one interviewer would cover 15 households a day [(6 hours  $\times$  60 minutes)  $\div$  23 minutes]. A two-interviewer team will cover 30 households a day, which then becomes the cluster size.

To estimate the total duration of your survey, you should also allow for travel time from town to town.

#### **EXAMPLE:**

In chapter 2, we estimated that 20 interviewers would cover 4,500 households in 15 working days but we allowed a further five days for travel from town to town and for unpredicted delays.

You can calculate the number of interviewers needed using this formula:

$$\text{Number of interviewers} = \frac{\text{sample size}}{\text{number of days available} \times \text{households per interviewer-day}}$$

#### **EXAMPLE:**

If the sample size is 4,800 and you want the work done in 20 days, and if each interviewer can do 15 interviews a day, the required number of interviewers will be  $4,800 \div (20 \times 15) = 16$ .

*Do not use more than 15–20 interviewers per region.* It will be difficult to ensure good quality training and supervision for a larger number. Furthermore, if the survey will be done by different teams in each region or district, *make sure that the training for each team is the same.* It is best to use the same trainers and training materials for *all* the survey field training.

Once you have decided how many interviewers are required, work out the team compositions. Each team will need one supervisor, one driver (unless public transportation is used) and interviewers. A common team composition includes six persons: supervisor, driver and four interviewers who may work in pairs, visiting alternate houses. This arrangement assumes that the vehicles will be large enough to carry six persons plus their equipment, questionnaires and personal luggage.

Team morale is easier to maintain when the interviewers work in pairs (see chapter 6). Working in pairs also contributes to quality control because one interviewer will always be close to another and they will make decisions together on selection of households and other interviewing matters. It is also useful to rotate team members to avoid monotony.

### **Arranging Transportation, Accommodations and Security**

Transportation may be provided by government offices or arranged privately—for example, by renting cars. When using government vehicles, ensure that they are well maintained and that there will be no conflicting demands for the vehicles during the field-work period. Allow funds in the study budget for fuel, maintenance and eventual repairs, or ensure that they will be covered by government funds. Estimate fuel needs by calculating the typical distances to be travelled from town to town and within each selected area.

It is often possible to arrange for the team's accommodations with local communities: teams may often sleep in church buildings, army quarters or even the mayor's house. If private accommodations are arranged, make sure that the interviewers get a daily allowance that is sufficient to cover their costs. Meals may also be arranged with the local authorities, as mentioned above. Many places have no commercial restaurants, so meal arrangements will have to be made in advance.

Security issues are also important. Field work may take place in urban slums or in rural areas where there may be security problems. Local guides are often useful in indicating and avoiding

---

---

**When in doubt, be conservative. If you are not sure whether each interviewer can do 15 or 20 interviews a day, choose the smallest number. Allow plenty of time for travel and for rest because field work can be very tiring.**

---

---

---

---

**Make plenty of allowance for extra mileage, since the actual distances to be travelled are often underestimated.**

---

---

security risks. These issues must be considered in advance.

Finally, careful arrangements should be made for paying the field workers and supervisors, as well as providing them with "pocket money" for meals, accommodations and other unexpected expenses. Timely payment is essential for maintaining the team's morale.

---



---

**Adequate arrangements for transportation, accommodations, meals and security are essential not only for ensuring high-quality and timely data collection but also for the psychological well-being of the interviewing team.**

---



---

### **Obtaining and Preparing Copies of Local Maps**

Before the field work begins, you should obtain copies of the maps indicating the large areas (states, provinces, districts, towns, etc.) as well as the small areas (villages, census enumeration areas, etc.) in which the survey will be conducted. These may be available from the Census Bureau or another government office. Army maps are often very useful, if it is possible to gain access to them. Make sufficient copies of all maps in advance.

## **PREPARING THE QUESTIONNAIRE**

The questionnaires you need for conducting your survey are given in chapter 3. Before training begins, you will need to translate the questionnaire and the instructions for the interviewers into the local language. As explained in chapter 3, do not expect the interviewers to translate the questions as they ask them. Different interpretations of the questions will make the data useless.

The questionnaire should be translated by one person and then another translator must independently translate the questionnaire back into the original language. The two versions can then be compared. Discuss any words which seem to be ambiguous or confusing, and agree on the correct translation.

---



---

**All the interviewers must ask the questions in the same way.**

---



---

When more than one local language exists in the area to be surveyed, use this translation procedure for all the questions and the instructions, for each language that will be used. Remember to give the translators very clear definitions of all the terms used in the questions. The correct definitions are given in the Instructions for Interviewers (Appendix 1).

#### **📎 EXAMPLE:**

Make sure the order of the questions is not changed during the translation process. Take particular care over the translation of phrases such as "continue feeding or eating" and "since this time yesterday." Be careful, as well, when referring to answers from previous questions (e.g., "during this last episode of diarrhoea").



Finally, keep the questions simple and easy to understand. Useful suggestions for phrasing the questions are given in chapter 3 (Box 3.4).

### Pretesting the Questionnaire

You must pretest the translated questionnaire in the field. The pretest should identify potential problem areas, such as dates of birth or of vaccinations, unanticipated interpretations and cultural objections to the questions. Apply the pretest to respondents similar to those who will be interviewed during the survey. The survey coordinator should do the pretest with the help of one or two future supervisors or interviewers. Do not make all the copies of the questionnaire for the survey until *after* you have pretested it.

Specifically, the pretest should answer the following questions:

- Are respondents willing to answer questions in the way you have asked them?
- Are any of the questions particularly difficult to answer or do they address sensitive issues?
- Are the questions well understood by the respondents?
- Can the interviewers follow the instructions easily, or do they misinterpret them?
- Is the questionnaire designed with adequate space and is the coding of answers clear?
- Is it necessary to create new codes for common answers which were not included in the original questionnaire?
- How long does an interview take? (This will help you to decide how many interviewers you will need.)

☞ Note that during the pretest the interviewers are still learning the questionnaire, so the time spent per interview is longer than it will be in the field after they become more experienced.

Discuss the results of the pretest with experienced colleagues and with the interviewers. Make any changes necessary to the Instructions to Interviewers (Appendix 1). If the pretest reveals that respondents refuse to answer the questions in the form in which they are given in the questionnaire, consult an experienced survey worker in the country in which you are working. These experts can help you to decide whether it is advisable to make changes to the questionnaire. If a significant number of respondents refuse to answer the questions, the survey will not be worth doing.

Once the questionnaire has been translated and pretested, you will need to make copies of it. When making copies, remember the following:

- Ensure that space for the identification codes appears on each page if your questionnaire is more than one page long. This will prevent any part of a household's questionnaire being lost.
- Use good quality paper. This will help you to write clearly and will prevent the questionnaires from tearing.
- Do not change the layout of the questionnaire. In particular, do not try to squeeze too many questions in a page. A good layout helps to reduce interviewer error in the field. If you use the layout given for the questionnaires, data can then be entered directly into the computer. This saves time and effort.
- Print more copies than you need. There will always be some wastage, and extra copies are needed for training. Allow a separate questionnaire for each household in your sample.

☞ Check that your questionnaire contains sections for:

- interview date
- interviewer's name (or number)
- a unique identifying number for cluster (e.g., 01), household (e.g., 02), mother (e.g., 1-0) and eligible child (e.g., 1-1)
- an introductory paragraph written for your survey, explaining the purpose of the survey, asking for permission to do the interview and stating that the information obtained during the survey is confidential
- introductory paragraphs to the different modules—for example: "I would now like to ask you a few questions about diarrhoea."

*The interviewer guide should be translated with the same care as the questionnaire.* Pretest it by giving it to potential interviewers, having them read it and discussing it with them to identify any instructions which are unclear.

## SELECTING THE FIELD WORKERS

The quality of the information obtained from a survey depends on the quality of the work done in the field. Good survey organization and thorough field work are vital.

A team of interviewers and their supervisors will do the field work. A detailed description

of the interviewers' and supervisors' tasks is given in chapter 6, but a brief job description for each follows to help you identify potential candidates.

The *field supervisor's job* is to:

- identify the clusters to be surveyed
- supervise three to five interviewers as they perform the survey
- ensure that the interviewers follow instructions
- answer interviewers' questions as they arise
- control the data quality by checking for errors during the interviewing, by checking that forms are completed fully and correctly and by checking that all the respondents are answering the questions
- identify problems and retrain interviewers who are doing their job incorrectly

The *interviewer's job* is to:

- identify the specific households to be surveyed
- gain the consent of respondents to be interviewed
- conduct interviews using the standard questionnaire
- maintain standard procedures in conducting the interviews and recording the answers.

You must select the interviewers and supervisors for their ability and motivation to perform these tasks. Supervisors must understand the importance of adhering to survey instructions, and be capable of ensuring that interviewers follow instructions.

The interviewers and supervisors should be:

- intelligent and literate—educated to secondary school level or more
- willing to follow instructions precisely and accurately
- polite and able to establish a good relationship with the respondents
- fluent in the language of the respondents
- employed by an institution or organization that will use their training for this survey in future; this training will then be part of capacity-building for future surveys

Previous survey experience is not necessarily a positive factor. While participation in well-conducted surveys is surely an advantage, previous involvement in poorly planned and implemented surveys can lead to bad interviewing habits which may be hard to correct.

 **EXAMPLE:**

In many countries the standards of so-called market research are very poor. Interviewers with previous experience in these surveys may actually require more training than totally inexperienced candidates.

You should also try to avoid over-qualified interviewers, who may follow their own agenda and stray from the precise technique developed for conducting the survey.

 **EXAMPLE:**

In some countries medical doctors were used for data collection, often with disastrous consequences due to their inability or unwillingness to follow the questionnaire instructions precisely and their tendency to make medical diagnoses during the interview.

Use female interviewers if possible, and ensure that the age of the interviewers is adequate for the information you want. In some societies women may be reluctant to provide answers on sensitive issues such as pregnancy outcomes or breastfeeding to male interviewers or to interviewers who seem too young.

In addition to the above qualifications, supervisors should have previous field experience as interviewers in well-conducted surveys.

Always select more potential interviewers than you will need. Train all of them and select the required number at the end of the course. This will guarantee that only the best

---

---

**Always select more potential interviewers  
than you will need.**

---

---

field workers will be involved in the study, and will also provide a few additional interviewers in case you need replacements. Provide those who were excluded with a training certificate.

## **CHOOSING AND PREPARING THE EQUIPMENT**

Equipment must be purchased well in advance of the survey. Box 5.1 lists some of the main items of equipment that must be acquired. (Cluster forms and maps, which are not listed, are discussed elsewhere.)

**Box 5.1**  
**SAMPLE LIST OF EQUIPMENT FOR FIELD WORK**

Notebooks for the supervisors	Pencil sharpeners
Clipboards	Envelopes for filing questionnaires
Backpacks or other types of bags	Paper clips and staplers
Blank growth charts	Weighing scales and accessories
Pencils and erasers	Length/height meters
	Salt iodine testing kits

☞ In addition to the above-mentioned equipment, field workers should also carry letters of introduction to the families, preferably on official letterhead, and identification cards with photographs.

### **Weighing Scales**

The scale must be portable, resistant and capable of weighing children up to 25 kg. Ideally, each interviewer will have her own scale. Interviewers who work in pairs can share one scale, but this will slow down data collection. A few extra scales should be ordered in case of equipment breakdown, loss or theft.

Two basic scale models are available: electronic scales and hanging scales.

The **UNICEF electronic scale** (available beginning December 1994) is a floor scale for weighing children as well as adults (capacity 150 kg). It has a precision of 100 g and a digital display. The child may be weighed directly, if possible. Alternatively, if the child is frightened, the mother can first be weighed alone and then weighed while holding the child in her arms, and the scale will automatically compute the child's weight by subtraction. Unlike hanging scales, there is no stress to the child and there are no trousers to wash (see next paragraph). No calibration is required. The scale itself weighs 4 kg and is powered by a battery with a life of 10 years. It costs approximately US\$90.

**Hanging scales** are widely used in many countries and are generally quite accurate. UNICEF distributes a Salter type spring scale (UNICEF catalogue no. 01 455 50) with a capacity of 25 kg and 100 g gradations. The use of this scale requires that the child be dressed in a set of plastic or nylon trousers before being weighed—a procedure that may make children anxious. (The same trousers can be used for more than one child, but they should be washed daily with water and soap. At least five

sets of trousers per scale should be available.) The interviewers will need several pairs of these special trousers if they are going to use hanging scales. A hook for hanging the scale from a door or a ceiling beam may also be necessary. The scale, which costs about US\$30, should be checked periodically with standard 5- or 10-kg weights.

### **Length/Height Boards**

Measuring length or height is optional in multiple-indicator surveys. Since children under two years will be measured lying down (length) while older children will be measured standing up (height), measuring boards should be adequate to both situations. As with scales, one measuring board per interviewer is desirable, but two interviewers working as a pair can share one board.

UNICEF distributes a model made out of wood (catalogue no. 01 145 00) adequate for children up to 130 cm, which is appropriate for the purposes of multiple-indicator surveys. It weighs 6 kg, measures 75 cm when collapsed and comes with a shoulder strap. It costs about US\$290. Alternatively, these boards may be manufactured locally at lower cost by skilled carpenters, but you should allocate plenty of time to this process since several adjustments in the early prototypes may be required.

### **Salt Iodization Kits**

Each field worker should carry a salt iodization kit. Each kit is sufficient for testing at least 100 samples of salt. Test kits for potassium iodate (catalogue no. 05 860 00), at a cost of US\$0.40 per kit, will be required for most countries. In a few countries—such as Ecuador and Colombia—where salt is fortified with potassium iodide, a different kit (catalogue no. 05 860 02) is required, which costs US\$0.60. These kits must be ordered well in advance of the planned start of the field work.

## **TRAINING THE FIELD WORKERS**

It is essential to have high-quality data. This will be possible only if you allow enough time to train the supervisors and interviewers thoroughly.

Before training starts, you should work out the field procedures to be followed during the survey. This means planning for supervisors to check completed forms and to fill out cluster control forms, for collecting completed

forms from the supervisors, for making transportation arrangements for teams and for making payments to field workers.

---

---

**Decide which interviewers and supervisors are qualified to go on and dismiss those who cannot perform the tasks adequately.**

---

---

☞ Remember to:

- plan ahead for the training course
- prepare interviewer guides (see chapter 3)
- make sure adequate space is available
- provide facilities for drinks and snacks (a good working atmosphere during the training course can help to motivate interviewers to perform well in the field)
- use audiovisual aids, such as overhead projection, during the training

Before you train the field workers you should also:

- translate and pretest the questionnaire, the instructions for filling in the questionnaire and the field procedures
- identify typical field locations for practicing household selection and interviews.

Table 5.1 provides an example of a five-day training course for both interviewers and supervisors. The latter will also need additional training (Table 5.2).

**Table 5.1. Example of a five-day training course for interviewers and supervisors**

<b>Day 1</b>	<p><b>Explain thoroughly the purpose of the survey.</b></p> <p>Introduce all team members, participants from the Ministry of Health and other organizations.</p> <p>Outline the whole survey procedure.</p> <p>Motivate the field workers by explaining the importance of the data to be collected and what will be done with it.</p> <p>Explain the administrative arrangements for the work.</p> <p>Give details of the working hours and pay, the survey schedule, transportation arrangements and everyday procedures.</p>
--------------	--

*(table continues on next page)*

---

**(Table 5.1, continued)**

---

**Days 2–3****Discuss the survey procedures and questionnaire.**

Conduct a question-by-question discussion of the questionnaire.

Explain and discuss each question. There should be no unfamiliar terms. Give each field worker written instructions to take to the field.

Discuss interviewing technique. Explain how to gain the confidence of the respondent, how to avoid inducing answers, the importance of completing each assigned interview and of following standard procedures. Emphasize that the interviewers must ask the questions *exactly* as they are worded on the questionnaire.

Standardize anthropometric techniques. Spend at least half a day in a place with many small children (day care center or nursery).

Do a demonstration interview.

Practice recording data, managing forms, making preliminary tabulations.

Have role-playing interviews, where trainees interview each other. Use questionnaires completed in the pretest as examples.

Tape the practice sessions if possible, and provide constructive criticism of the different interviewers. Hold more demonstration interviews as the training proceeds.

---

**Days 4–5****Conduct a field exercise and have further discussion of interviewing.**

Practice reading maps.

Discuss how to handle empty buildings and refusals.

Organize practice in the field. Each trainee should complete at least five practice interviews in the field. Observe all the interviewers' practice sessions and provide them with feedback.

Discuss the problem of the interviewer influencing the respondents' answers and other interviewer mistakes. Agree upon solutions to these problems.

Go over field practice questionnaires with individuals who have particular problems, and discuss problems as a group.

Ask the participants to share their ideas and suggestions for dealing with difficulties.

---



**Table 5.2. Example of a two-day additional training course for supervisors**

<b>Day 1</b>	<p><b>Household selection and map reading.</b>          Explain the procedures to be followed, and the importance of random selection of households.          Provide practice and time for discussion.</p> <p>(If sketch-mapping will be used for segmentation of the small area, as described in chapter 6, then at least two additional days for training in the field will be required.)</p>
<b>Day 2</b>	<p><b>Quality control.</b>          Explain the need to monitor interviews and check interview quality on the spot.          Discuss how to deal with interviewer errors.</p> <p>Explain what to do with the completed questionnaire, and how to deal with unanticipated problems.</p> <p>Emphasize that the supervisor should keep field notes, and go through what should be recorded in these notes.</p> <p>Discuss the survey schedule and the need for liaison with the survey coordinator.</p>

### Briefing the Drivers

All persons involved in the survey should be briefed about its main purposes and the main methodological guidelines. Drivers who will work along with the team throughout the whole survey are a group in need of special attention. Drivers often fail to understand random sampling and may even refuse to take secondary roads or paths to reach scattered households. This preference for certain roads is known as "main road bias." Another common problem is that drivers often interfere in the interviews and prompt or induce certain answers from the mothers. A special session at the beginning of the field work may help prevent these problems.

### CARRYING OUT THE PILOT STUDY

The pilot study is the final rehearsal for the survey. It should be carried out soon after finishing the training period, but at least a few days before beginning the actual field work. This will allow time for correcting any problems detected during the pilot study.

**A properly conducted pilot study will identify the major problems with the survey methodology and help prevent them during the data collection phase.**

The pilot study should cover both urban and rural areas. These areas should be selected to be representative of the situations the interviewers might face during the survey. The pilot study should last for three to five days and include the daily routine shown in Box 5.2.

**Box 5.2**  
**DAILY ROUTINE FOR THE PILOT STUDY**

- ✓ Briefing at headquarters
- ✓ Transportation to field sites
  - ✓ Locating clusters
- ✓ Contacting local authorities and introducing yourself
- ✓ Selecting the households
- ✓ Interviewing and measuring
  - ✓ Recording data
- ✓ Managing data and tabulating preliminary results

The pilot study should be seen as an extension of the training program. Close supervision of the interviewers during this phase is essential.

### **SETTING UP COMPUTERS AND HIRING DATA PROCESSING STAFF**

You must obtain the services of a computer programmer with experience in using the EPI INFO software package. (See chapter 7 for more information on EPI INFO.) You will need this person for at least two to three months. Copies of EPI INFO are available from the Planning and Coordination Office in New York.

Standard EPI INFO programs for entering, cleaning and analysing the data collected are described in chapter 7. However, these programs may have to be adapted locally, particularly if any changes are made to the questionnaire modules (chapter 3). Before the main survey begins, make sure that the programs have been properly adapted and are functional.

- ☞ Use the questionnaires of the pilot study for testing the data entry and analysis programs. Check the programs for the production of tables. Sort out any problems and make any corrections that may be necessary.

You will also need data-entry staff. Depending on the size of your survey, two or more data clerks will have to be recruited and trained by the computer programmer in using the EPI INFO software. This training should require no more than two days.

Arrange for the necessary office equipment, including computers, printers, diskettes and paper, and make sure the power supply is adequate. At least two computers are required for the survey work.

***Remember—unless all arrangements for data entry and analysis are made before starting the field work, this process can lead to major delays in producing survey results.***

### **MAKING ARRANGEMENTS FOR RETURNING THE QUESTIONNAIRES TO HEADQUARTERS**

Instruct supervisors on the procedure for returning completed questionnaires to the data-processing headquarters.

If possible, the questionnaires should be returned daily so that data can be processed quickly. In the early stages of the survey this will also enable you to check for any systematic problems which may still be occurring in the field. To prevent losing the questionnaires, arrange for supervisors to photocopy completed questionnaires before sending them long distances, if at all possible. If this is not possible, at least arrange for small batches of questionnaires to be dispatched at any given time so that any losses will be minimized.

**The prompt return of the questionnaires to headquarters contributes to quality control and allows early data entry.**

#### **EXAMPLE:**

In one survey, an arrangement was worked out with the local bus companies that drove to each settlement in the country. No batches were lost, and no more than a few days were lost sending the questionnaires from the field to the data processing center.

### **ETHICAL CONSIDERATIONS**

Household surveys typically raise a number of ethical questions, particularly surveys that pertain to the health of children and other household members. Such questions relate to individual rights to privacy, the need for informed consent, and responsibilities that arise upon uncovering potential health problems in a survey. It is important to consider such dictums as those enumerated in Box 5.3 during the early stages of planning a survey.

**Box 5.3****ETHICAL ASPECTS OF CONDUCTING A SURVEY**

**Ethical approval:** The survey must abide by the laws of the country. If approval by an ethical committee is required, this should be requested at an early stage to prevent delays.

**Confidentiality:** All information provided to the interviewers is strictly confidential. Records should be securely stored. Computerized records should not include any names that might be used to identify the families, unless this is strictly necessary (for example, if follow-up visits are being planned).

**Informed consent:** Mothers and/or heads of families should be informed about the contents of the interviews and measurements to be carried out. They must understand the procedures and give their full approval. In some countries, written consent may be required.

**Feedback to the families:** Families have freely donated their time to the survey and are entitled to some feedback. Any important conditions discovered during the interview should be reported to the parents. For example, mothers should be advised when their children's vaccinations are overdue, when the child is malnourished, or when non-iodized salt is being used. In some countries, field workers carry ORS packets or plastic spoons for preparing sugar-salt solution to distribute to children with diarrhea who are not being treated properly.

**Feedback to community:** Before starting the survey, the coordinators should plan what type of feedback will be given to the communities. In most cases, the number of interviews per community will be too small for statistical validity, but even some general feedback is often appreciated by local authorities (for example, that 9 of the 40 children examined were found to be malnourished). If possible, this type of feedback should be given before the team departs to a new community.

## CHAPTER VI

# CONDUCTING THE FIELD WORK

*This chapter should be read by technical resource persons and survey coordinators. Supervisors and interviewers should read the sections that are relevant to them. The chapter will teach you:*

- ✓ How to map and locate the households.
- ✓ How to handle some special situations in the field.
- ✓ What interviewers should do in the field.
- ✓ What supervisors should do.

### HOW TO MAP AND LOCATE THE HOUSEHOLDS

In chapter 4, you were asked to choose a sampling procedure for households within small areas such as villages or census enumeration areas. We now give detailed guidelines on how to select households using three alternative methods: segmentation, starting point based on a boundary map or EPI-style random walk. Read carefully the option corresponding to your sampling design.

In this chapter, the following definitions are used:

- A *household* is a group of persons who eat and live together.
- A *dwelling* is a building or residential unit. It may include one or more households, as in the case of compounds or apartment buildings.

### USING SEGMENTATION TO IDENTIFY THE CLUSTERS

Segmentation must be done by a supervisor, usually one or more days before the field work. Follow the steps below for each small area selected:

1. **Compute the number of segments to be created** by dividing the total number of households in the small area—according to the last census—by the desired cluster size. Table 6.1 gives the number of segments to create in an area when the cluster size is equal to 40.

**EXAMPLE:**

Suppose the census recorded 250 households in the village and the cluster size is 40. Dividing 250 by 40 gives 6.25. Round it to the nearest whole number, in this case 6. Exactly six segments are to be made in this village. You would have reached the same result by using Table 6.1.

**Table 6.1. Determining the number of segments to create in a given small area (cluster size = 40)**

1–60	1
61–100	2
101–140	3
141–180	4
181–220	5
221–260	6
261–300	7
301–340	8
341–380	9
381–420	10

*etc.*

In some cases the area may be very populated—more than 1,000 households. You may divide these first into four quadrants with approximately the same area and then select one of the quadrants at random before carrying out the segmentation. In this case, the census count must be divided by 4.0 before calculating the number of segments in a quadrant. A note must be made of these special cases so that the proper survey probabilities and weights can be calculated later at the analysis stage.

**EXAMPLE:**

A village with 1,400 households may be divided into four quadrants of about 350 households each. According to Table 6.1, above, nine segments would be created in the chosen quadrant and one of these segments selected at random.

**2. To prepare the sketch map, go to the field with the available boundary map and locate the outer boundaries of the small area.** If a map is not available, get help from a local person and draw a map marking these boundaries, identifying the names of roads, lanes, streets and showing physical boundaries such as streams, rivers, and so forth.

3. **Draw internal markers**, which will help identify locations and aid in establishing a path of travel. These would include internal streets, paths, streams and so forth.

4. **Draw a small square representing each dwelling unit in its appropriate location on the map.** For help in later locating the households, it is also useful to mark other prominent buildings—schools, churches, mosques, and so forth.

☞ Note that the sketching should be a "quick count" operation for locating dwellings. It does not require knocking on doors to inquire about households or the names of occupants. An exception is multi-dwelling buildings that appear to include several households—for example, in the case of large compounds or apartment buildings, where you should ask about the number of households and record it on your map.

5. **Count the number of dwellings in the small area** and divide it by the desired number of segments calculated in step 1 above. This will give you the actual number of households in each segment.

**For purposes of making the segments, absolute precision in the count of households is not necessary. A quick count of dwellings can be used instead of counting households.**

6. **Based on the sketch map, divide the area into segments with roughly the same number of dwellings.** This means that the size of the segments in terms of area is likely to vary considerably—densely populated areas will have geographically small segments and low-density segments will be large. Give each segment a number.

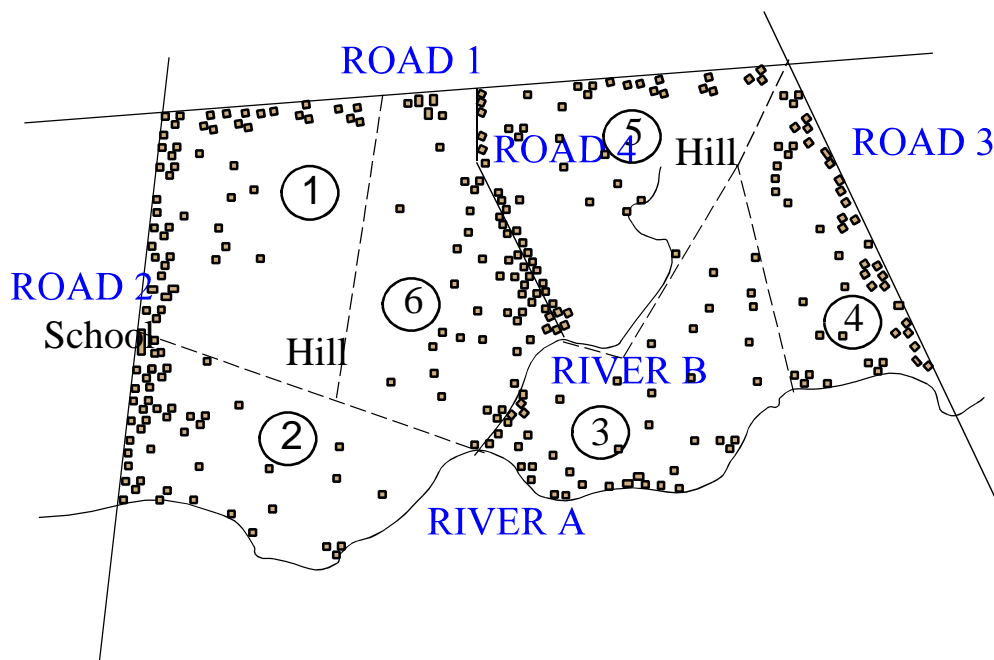
📎 **EXAMPLE:**

You had calculated that you needed six segments and, after mapping, your area was found to include 310 dwellings. You should then divide it into six segments of 51 dwellings each.

☞ It is essential to create segments in such a way that segment boundaries can be easily identified. You may have to relax the criterion of equal-sized segments in order to achieve this. The segment will be visited later, perhaps by some other field worker, so the boundaries must be readily identifiable on the sketch map and on the ground.

**The number of segments should be calculated on the basis of the last census count even if the population within the small area has changed since the census. This will ensure that changes in population are automatically adjusted for.**

**Figure 6.1. Map of hypothetical rural area showing creation of six segments with approximately 50 dwellings each**



7. After segmentation has been completed, the next step is to select one and only one segment at random in each of the small areas. Using a random number table or writing the numbers in pieces of paper, choose a random number between 1 and the number of segments created in the area to identify which segment is to be selected.

**EXAMPLE:**

If the small area has been divided into six segments, choose a random number between 1 and 6 to determine which is selected. Do this independently for each selected small area.

If possible, the person selecting the segment should not be the same one who created the segmentation in the first place, in order to prevent the possibility of inadvertent bias.




**Box 6.1**  
**USING SEGMENTATION IN URBAN AREAS**

Segmenting urban areas is easier than segmenting rural areas. Cities and towns are usually organized into blocks or some similar units. When using census enumeration areas, maps are usually available showing streets and blocks. If unavailable, these maps can be easily drawn. A quick drive through the area will let you find out whether the number of dwellings appears to be about the same in every block. If so, the segmentation may be based on the blocks.

 **EXAMPLE:**

Let us suppose that your urban census tract includes 17 blocks and that you require seven segments. Divide 17 by 7 to obtain 2.4 blocks per segment. Using the map, divide the area into seven segments with two and a half blocks each (the last will have to include three blocks to complete 17) and choose one of these at random.

If the number of dwellings in the blocks are variable, proceed as for the rural areas, counting the total number of dwellings in the area and later dividing them into the required number of segments.

-  Although when sketch-mapping rural areas you do not have to worry about separating dwellings from households, in urban areas you should ask how many households are present in tall buildings (for example, those with more than two floors).

## **THE RANDOM WALK METHOD**

The random walk method includes two separate steps. The first is to choose a starting point and the second is to select the households from that point onward. According to the flowchart in chapter 4, whenever boundary maps are available they should be used for selecting the starting point. If these maps are not available, then the EPI-like method should be used. In both cases, the random walk method should be used for selecting subsequent households from the starting point. These methods are described below.

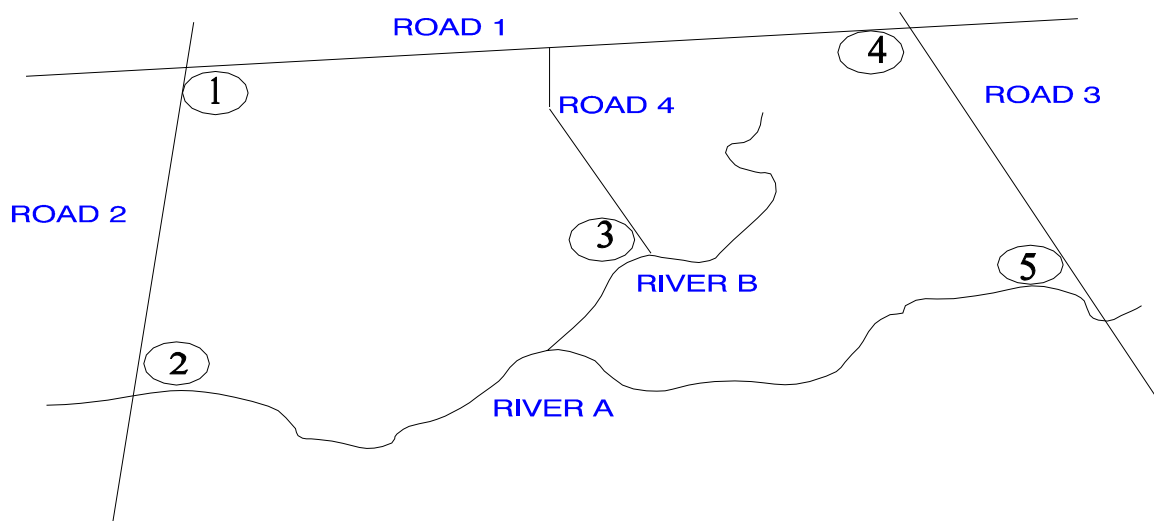
### **Selecting the Starting Point on the Boundary Map**

Using a map of the outline of the small area (Figure 6.2), follow the steps below:

1. Allocate a number of starting points (usually four to seven), spread out on the map.

Typically, these points are put on easy-to-locate spots along the boundaries of the area, but at least one or two additional points are placed in the middle of the geographical area. Ensure that these points are linked to a landmark so that the field workers will find them without difficulty.

**Figure 6.2. Map of hypothetical rural area showing starting points**



2. Choose one of these numbers at random—for example, by writing them on pieces of paper and drawing one.

☞ It is not necessary to go to the field before completing the above steps. You do not have to know where the households are located when the starting points are allocated.

It is possible that with this sampling scheme some households will have twice the chance of being selected (for example, those exactly in the middle of two of the numbers), while others will have a zero probability of selection (for example, those far from all numbers). This sampling scheme, therefore, does not produce a strict probability sample.

Also, it may be argued that this method gives a greater chance of selecting an isolated household as the starting point. This is more likely to be true if there are great differences in household concentrations within the area, and less likely if households are more evenly scattered. One should note, however, that having selected this starting household one would proceed within

the area to find the subsequent ones. Any concentrations of households such as small villages located near that starting point may be fully included in the sample.

Despite these shortcomings, this method may result in a sample without important biases, being particularly applicable to scattered rural populations in low-density areas, where enumeration or segmentation would be costly and lengthy.

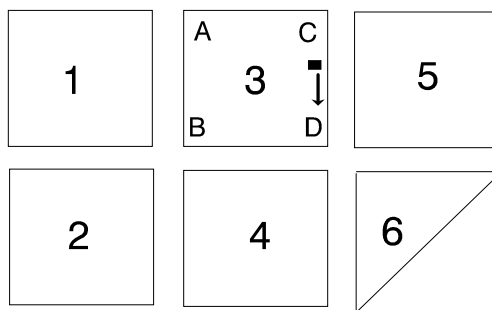
**Box 6.2**  
**USING THE STARTING POINT METHOD**  
**IN URBAN AREAS**

The starting point method can be easily adapted to urban areas. After numbering the blocks, select one of them at random. The corners of the chosen blocks are assigned letters, one of which is also selected at random. Moving in a clockwise direction from that corner, all houses up to the next corner are numbered and one of these is also randomly chosen. That will be the first household to be interviewed. From there onward, every third or fifth household, for example, would be included in the sample. If the selected block does not include sufficient households, the next block (in numerical order) would be selected for completing the cluster.

**EXAMPLE:**

Figure 6.3 shows a map of an urban enumeration area composed of six blocks. Block 3 was selected at random, and then corner C was selected in that block. One house on that side of the block was chosen and subsequent houses were selected in a clockwise direction.

**Figure 6.3. Example of the random starting point method in an urban area**



### Using the EPI Method for Selecting the Starting Point

This option is particularly useful if you have sampled small areas (for example, villages or rural communities) but there are no boundary maps that would allow you to use segmentation or to choose a starting point on the map.

Field workers should follow these steps for selecting the households:

1. Go to some central location within the community. This could be a market, a church, a health facility or the junction between two roads. If possible, find a local guide who can accompany you. He/she can introduce you to the families and help you find the households and the community boundaries.
2. Select a direction at random by spinning a bottle. Move in a straight line in this direction. Count all the houses you pass until you reach the edge of the community. If possible, number each house with chalk as you go.
3. Pick one of these numbered houses at random to form the starting point for the survey.

Like the previous method of selecting a starting point from a boundary map, the EPI method does not produce a strict probability sample. It also has an additional disadvantage of relying on vague definitions of "village boundaries." Under some circumstances, however, when boundary maps are not available, it is an acceptable method.

### Using the Random Walk Method for Selecting Households

Whether the starting point was selected from the boundary map or through the EPI procedure, the random walk method may be used for selecting the next households.

Box 4.6 in chapter 4 described the original random walk method. The guidelines below have been slightly modified in an attempt to overcome some of the problems with the standard method. The main modifications are:

- Interviewing should stop when the required number of households (not children) is reached.

 **EXAMPLE:**

If the cluster size is 30 households, field work should be interrupted when this number is reached, even if only a few children were found in that particular cluster. Other clusters will have more children than expected so that in the end numbers will balance out.

- Special caution should be taken to avoid bias when selecting households.

**EXAMPLE:**

Households should be selected in such a way that those far from the village center, or those distant from a main road, have the same chance of being sampled as more accessible households. Isolated families may have a different health status and it is important that they are not excluded from the survey.

- Further attempts should be made to interview residents who are not at home when initially visited.

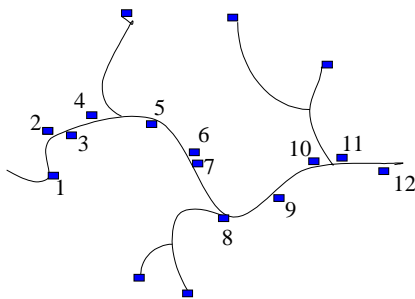
**EXAMPLE:**

Children whose parents are often away from home may differ in health status from the remaining children. Extra efforts, such as going back to the house at the end of the working day, are required for reaching these children and their parents. (This issue is discussed in greater detail later in the chapter.)

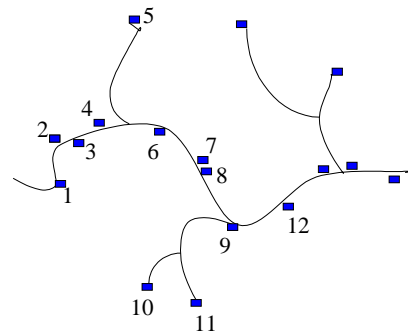
Follow these steps to select the households:

- The household nearest to the starting point should be the first to be included.
- Spin a bottle or flip a coin to decide in which direction you should move after that.
- Follow a road or a path to identify the next households. Figures 6.4A and 6.4B show that by choosing the next closest household you may leave out families living away from the main roads or paths. However, virtually every house will be connected to a road by an access path. By following these paths from the point where they join the main road (Figure 6.4B), you may avoid this type of bias.

**Figure 6.4A. WRONG: Selecting the nearest household will leave out the more isolated families.**



**Figure 6.4B. RIGHT: Following the access paths will also include the isolated households and avoid “main road bias.”**



- ☞ In scattered areas, the assistance of a local guide is essential for saving time. This person will help you follow the paths leading to households and avoid wasting time, following only those paths which lead to fields or empty houses.
- ☞ If you are uncertain about which way to move, use a random selection process to decide which household to interview.

📎 **EXAMPLE:**

If you are unsure about which of two households is the next to be interviewed, flip a coin to decide which household to go to.

4. Continue in this way until you have reached the required number of households (not children) in your cluster.

In urban areas, a procedure like that described for the starting point method may be used.

### *Using Different Sampling Schemes for Urban and Rural Areas*

You may use different schemes for sampling households in urban and rural areas. You may also use different sampling schemes in densely populated or less densely populated rural areas.

#### **Box 6.3**

##### **OPTIONS FOR IMPROVING THE RANDOM WALK METHOD**

You can improve your sample by spreading it out within the small area or community. Some suggestions are:

- ☞ Instead of selecting consecutive households, use systematic sampling. For example, interview one household and skip four; interview another one and skip another four; and so on.
- ☞ Divide the community or census tract into four sections (no mapping required) and select one-fourth of the required households from each quarter, using the same procedure for choosing the starting point within each section.

The above suggestions will improve your sampling, but you will need more time to complete the cluster since you will have to travel greater distances from house to house. In an area with low population density, spreading out the sample may not be feasible.

**Any method which achieves a random selection of households, preferably spread widely, is acceptable.**

 **EXAMPLE:**

You may interview every household in scattered rural areas and every third or fifth one in urban or dense rural areas (such as areas with many compounds).

## HOW TO HANDLE SPECIAL SITUATIONS IN THE FIELD

Some special situations in the field, described below, may require specific solutions:

 ***The small area cannot be reached.***

This is a rare event but in some cases—during the rainy season, for example—the area cannot be reached because of poor road conditions. When this happens, survey the closest area to the original one that can be reached. All replacements should be recorded in the field notebook and mentioned in the final survey report. Note that replacing the cluster that cannot be reached with another area does not completely solve the problem, since children living in areas of difficult access are likely to also differ in health status and in the utilization of services.

 ***The small area includes fewer households than the required cluster size.***

Survey all the households in the area. Then move to the area that is closest to the last household you did in the original area. Survey as many households in this second area as you need to complete the cluster.

 ***Separate households are difficult to identify.***

If separate households are difficult to identify—for example, where there are a few related families living in the same dwelling and sharing some but not all meals—treat the entire dwelling as one household and interview all the women and children within that dwelling.

 ***The household is empty.***

If you call or knock a few times and there is no reply, ask the neighbors whether the house is inhabited. If it is not occupied, and you are working with a fixed-size cluster (for example, 40 households), select an additional household at the end of the day to compensate for the one which was empty. If you are working with a variable number of households per cluster—as when using segmentation—just go on to interview all the other households in the segment.

If the neighbors tell you that the house is inhabited, come back at the end of the working day to see if the residents have returned. If they are still away, try to come back on another day. This may not be feasible in rural areas where the whole team is moving from place to place, but in urban areas it is often possible.

☞ ***The residents refuse to be interviewed.***

Refusals are often a transitory reaction. Therefore, never accept a refusal as definitive. Ask one of your team colleagues to visit the house later in the day and try again. If another interviewer still does not succeed, ask the supervisor to try at the end of the day. Do not give up until three attempts have been made. If you still fail to interview, make a note of the refusal on the cluster sheet, but do not replace the house. Report the refusal to the supervisor.

Give higher priority to returning to clusters where there were several losses than to those with one or two losses. This is because residents of the former are more likely to be atypical—for example, they may be of higher socioeconomic status or many mothers may be out at work—and their exclusion could bias the sample.

**Bias may result if losses and refusals are common. Start to worry if you fail to interview more than 5–10 per cent of your intended sample.**

Houses where residents are absent or refuse to be interviewed should be carefully recorded and counted as losses in the final report. Children and adults living in these houses may well be different from those interviewed, so that replacing them with another (compliant) household does not solve the problem.

## **WHAT EACH INTERVIEWER SHOULD DO**

The daily routine of the interviewer should include the following:

- Read the cluster map and locate the segment or the starting household.
- Follow the instructions, to select the next households in the cluster.
- Fill in a new cluster form for each cluster.
- Introduce him/herself to the household members and ask permission to carry out the interview.
- Apply the questionnaire to all members of the target population in the household.
- If applicable, carry out the additional procedures for data collection, including anthropometry, testing salt iodine and observing water and sanitation facilities.



- Fill in the responses to the questionnaire in a neat and legible form.
- When in doubt about the coding of an answer, write it down in full at the margin of the questionnaire and check the correct coding with the supervisor at the end of the day.

### **WHAT EACH SUPERVISOR SHOULD DO**

The supervisors will have many tasks during the survey. Tasks related to sampling have been described above, and will vary according to the method used. In addition to these tasks, all supervisors will have to perform the following tasks on a daily basis:

- Contact the local authorities or leaders in every community and, if required, hire local guides.
- Supply the interviewers with questionnaires and other field work materials.
- Assign clusters or households to the interviewers.
- Observe about 5 per cent (that is, one in every 20) of each field worker's interviews and correct any mistakes that may be noticed. For example:
  - Check that the interviewers are stating the questions exactly as instructed.
  - Explain questions and answers that are not being properly interpreted.
  - Check that the interviewers are not inducing particular responses
- Repeat a few of the interviews and measurements separately from the interviewer, particularly in the early phase of the survey. Check that the answers are consistent with those obtained by the interviewer and, if not, take corrective action.
- Review questionnaires as they are completed. For example:
  - Check that answers are legible.
  - Verify the consistency of answers (for example, the ORT module should not be answered for children without diarrhoea).
  - Investigate high levels of non-response (that is, if an interviewer reports many refusals or empty households).
  - Check totals—the number of children with completed health questionnaires should equal the total number of children under five years of age living in the household.
- When a return visit is necessary, note call-back times and household position or instruct interviewers about how this should be done.

- Check that interviewers are not replacing households which are difficult to contact with other households.
- Change interviewing times if necessary to prevent the need for too many return visits.
- Answer questions, resolve problems, give feedback to the interviewing team on the progress of the survey.
- Keep the team on schedule. Allocate assignments to interviewers and re-deploy staff when necessary.
- With the team, make basic tabulations and feed back information to the community, as decided by the survey managers.
- Carry a notebook at all times to keep a record of:
  - sampling procedure (details on the clusters selected)
  - how households were allocated to the interviewers
  - the outcome of each interview (see cluster control sheet), transcribed from the interviewer's records
  - any difficulties in the field (also transcribed from interviewers' notes)
  - a tabulation sheet for recording questionnaires as they are completed (see cluster control sheet)
  - ad-hoc decisions on how to code unexpected answers
  - the supervisor's general observations about the survey
- Collect all the completed questionnaires and keep them clean and safe, photocopy them if necessary, and send them back to headquarters.

***The number of tasks requested of the supervisor is large. If sketch mapping is also being done, it may be wise to have more than one supervisor per team.***

**Box 6.4****KEEPING UP THE MORALE OF THE TEAM**

Particularly after the first two to three weeks of field work, it is important to keep the morale of the team high. Interviewing becomes a routine and standards may be relaxed. Here are some suggestions for maintaining the team's spirits:

- If possible, avoid having the field workers away from their families for more than two consecutive weeks.
- Ensure that the field workers understand exactly how much and when they will get paid and avoid any delays in paying them. Ensure that money for expenses (for example, meal allowances) is paid before the expense is made.
- The interviewers should work in pairs whenever possible. They can carry out simultaneous interviews in neighboring houses. This increases their sense of security and they will be able to help each other make decisions about locating the houses, and so forth.
- Special attention must be paid to drivers. You should brief them and explain the need for random sampling and for reaching households that are far away from the good roads. They must also be instructed about the need to avoid interfering with the interviews. Otherwise, they may try to influence the sampling and interviewing procedure.
- Within the limitations imposed by the work load, interviewers should have time to rest at midday and at the end of each working day, as well as having at least one full day off at weekends. Otherwise they will become excessively tired and this will affect the quality of their work.

## CHAPTER VII

# PROCESSING THE DATA

*This chapter is for survey coordinators and technical resource persons. It gives guidance to help you:*

- ✓ Carry out data entry.
- ✓ Check and edit the data, and create a "clean" data file for analysis.
- ✓ Evaluate the quality of the survey data.
- ✓ Calculate estimates of the indicators.

### INTRODUCTION

Data processing should take place as quickly as possible. First results of the survey, the preliminary report, should appear quickly, within a few weeks after the field work ends. A full debriefing of the survey team, with a review of the initial results, contributes to a good final report. The enthusiasm and interest generated by the field work will evaporate quickly when the field work ends, so be sure to keep the leverage the survey results can give you by reporting them quickly.

As part of field procedures to reduce error, interviewers should check through the questionnaires when they complete an interview. Errors can often be corrected on the spot, without the need to re-interview. Field supervisors should also check through all questionnaires immediately after collecting them from interviewers. Are they complete? Are they clear? Are the "skip" instructions followed correctly? The survey manager should also check all questionnaires in the earliest stage of the field work, and continue to check samples of the questionnaires throughout the survey.

Data entry is best begun while interviewers are still in the field. This allows you to spot and correct mistakes that certain interviewers or teams may be making. Serious problems that may escape the field supervisor's notice can then be picked up quickly, in time to re-train field staff and correct serious errors. Field supervisors are responsible for checking the questionnaires for completeness and consistency, and for classifying any responses the interviewer could not code. Except for checking on the numbers of questionnaires returned and for completing identifying numbers on the forms, it should not be necessary for office staff to do any editing or coding of the questionnaires that are returned to the office.

The data-entry system must be carefully designed and data-entry staff must be well trained and motivated. The programs for data entry, checking and analysis should be set up and verified

using data from the pilot survey before the main field work phase begins. Any problems should be ironed out, and the programs modified if necessary, before the survey data start to come in from the field.

There are four main phases for the computerized data-processing of the questionnaires:

1. data entry
2. data checking and editing
3. frequency distributions
4. variable construction and tabulations

A variety of software is available for help in performing these functions. With this handbook, a complete program for data entry, checking and analysis to calculate all the indicators of the seven Mid-Decade Goals (that is, those that can be measured with a multiple-indicator survey) is supplied for use with the EPI INFO (Version 6) package of programs. The programs are also listed in Appendix 3.

EPI INFO was developed by the Centers for Disease Control and Prevention in collaboration with WHO (available from USD, Inc., 2075 A West Park Place, Stone Mountain, GA 30087, USA) and can perform all four phases of the data processing. A copy of EPI INFO (Version 6) is included with this handbook.

The data processing supervisor should set up the data entry system in advance. Advice for doing so is found in Appendix 3.

## **ENTERING THE DATA INTO THE COMPUTERS AND PRODUCING "CLEAN" DATA FILES**

### **Data Entry**

As soon as data from one cluster arrive back at headquarters, data entry should begin. Data entry should be carried out in small batches (e.g., a cluster at a time).

Data from the questionnaire should be entered for each household in the order they are collected: the household data from the listing page, the data on all children on the child roster, the water and sanitation and salt iodization modules, education data for each child over school-entry age, tetanus toxoid module for each mother, care of acute respiratory illness (CARI) module for each mother and, last, child health modules for each child. Once all the data for one household are entered, the clerk moves on to enter the data for the next household.

---

---

**If any modifications have been made to the questionnaire, the data entry and analysis programs will have to be modified by a programmer with excellent knowledge of the EPI INFO program package.**

---

---

Note that missing information during the field work is recorded as 9's—for example, 9, 99, 999, and so on. Blank fields at data-entry time and inapplicable responses (e.g., for answers to optional questions or skipped questions) are stored as blank fields in the data base. As data for each household are entered, the program performs *range checks*. These ensure that no number is entered for an item that is out of the given range of responses for that item.

 **EXAMPLE:**

The date of interview should be recorded during the period of field work. A date of interview recorded as 01 11 93 is outside the range of dates allowed for this survey. If the data entry program has established the range for this variable, the computer will recognise that this date entry is invalid.

There are two approaches to entering data. Data can be entered into the computer twice, by two different data-entry clerks. The two files are then compared for inconsistencies, and other data checks performed. Another approach is to enter the data only once, but to perform a number of checks to prevent errors at the time the data are entered. Interactive data-checking can slow the data-entry process, but can avert many common data-entry errors made by clerks. Consistency checks are then run and the file corrected. For these multiple-indicator surveys, we advise the latter option, which is the approach that is used in the data-entry programs provided with this handbook.

**Box 7.1**

**PROS AND CONS OF DOUBLE ENTRY**

**Pros**

Data entry errors are virtually eliminated if done correctly. This option, followed by batch-error checking, may be cheaper and faster when professional data entry clerks are used.

**Cons**

Double entry doubles the cost of data entry. It complicates questionnaire flow in the office, which can also lead to errors.

A possible compromise is to enter some questionnaires twice as a way to spot-check the accuracy of data-entry clerks and identify systematic errors that are being made.

## Data Checking

Experience has shown that range errors are almost always data-entry errors and can be usefully checked for and corrected immediately at data-entry time. Consistency errors, however, must be resolved by returning to the household or by carefully examining the questionnaire. Consistency checking is therefore better carried out as a separate step, with errors reported on hard copy that can be used for marking in corrections.

The EPI INFO data-entry program supplied with this handbook automatically applies a large number of checks to these data, either alerting the data-entry clerk whenever an out-of-range response is entered, or reporting inconsistencies on the final pass through the data-checking program.

☞ Once the data are in headquarters, it is usually not possible to return to the households for correcting inconsistent answers. The study coordinator or a senior supervisor will have to make "best guesses" based on the existing information to correct the questionnaire. If that is not possible, the variable should be treated as "missing."

After data for an entire cluster are entered, the batch should be checked for consistency and edited. The program checks to ensure that all variables appropriate to the case are in the file (that is, the structure of the file is correct) and are within the correct range. The program performs internal consistency checks to ensure that the data have been recorded and entered properly. The data-entry instructions found in Appendix 3 discuss procedures to follow when errors occur. Reference to the original questionnaire at this stage can often help to sort out problems in the data.

### **EXAMPLE:**

Consistency checks verify that the following requirements are met:

- The date of birth for each child must precede the date of interview.
- Dates of birth for all children must be within the past 15 years.
- If the water source is "piped-in dwelling," then the distance should be "on premises."
- If a breastfed child received "ONLY breast milk," then no other items should be answered affirmatively.
- The date of first DPT should precede second DPT, which should precede third DPT.
- Anthropometric indices are not flagged as invalid due to errors in age- or weight-data entry.

It may be useful to look at the number of errors for each team or interviewer, so that further training may be carried out in the early stages of field work.

Once corrections have been noted for reported inconsistencies, the corresponding records are

updated on-line. The checking program should then be run again and the procedure repeated until all errors that can be corrected in the office have been rectified. After these checks are performed, the data are "clean" and the files are ready to be analysed.

## EVALUATING THE QUALITY OF THE SURVEY DATA

The next step is to execute *frequency distributions* of all the variables in your data file and the possible values they can take (see Table 7.1). This will allow you to check the data again for inappropriate entries and to note the number of missing values in the data.

**Table 7.1. Example of a frequency distribution**

Item number	Household water supply source	Distribution
1	Piped-in dwelling	146
2	Public tap	440
3	Tube well or borehole	102
4	Protected dug well	95
5	Unprotected dug well	660
6	Pond, river or stream	35
7	Tanker-truck, vendor	55
9	Other	15
.	Missing	6
<b>Total</b>		<b>1,554</b>

You may need to look back at the original questionnaire to check data entries. You may then need to correct the raw data file and perform this step again.

☞ Referring to the example in Table 7.1: Suppose you want to examine the 15 questionnaires with "other" water source. You will need to program the computer to select only those records that have a value of "9" (other) for the variable "water source." You must then tell the computer to list the identification number of the households with these values. You can use the identification number to go back to examine those questionnaires.

You should keep a record of how many cases in the file were incomplete, how many households were listed but not contacted and what errors were noted in the data. The per cent of eligible households which were not contacted should be examined and reported. A response rate



below 90–95 per cent means that the survey does not necessarily represent the entire population.

One way to check the quality of the data is to examine the internal consistency of the data collected. Consistency checks which were run at the time of data entry can be run again for the entire data set to determine the number of records for which inconsistencies remain. (See the example above on requirements for consistency.)

A high number of unresolved inconsistencies is an indicator of sloppy data collection (and sometimes falsification of data) and could damage the validity of results.

A second indicator of data quality is the per cent of responses which are "don't know" or "other." These percentages can be obtained by examining a frequency distribution for the items in the data set. A high per cent of "don't know" indicates that respondents did not understand the questions or that the information requested is too difficult to report. Results based on such variables are questionable. A high per cent of "other" responses indicates that the categories on the questionnaire do not capture the responses most commonly given.

Another indicator of data quality is how well they compare with other existing data. For example, the distribution of mothers and children by age and the ratio of boys to girls should be similar to the age distributions and sex ratios found in a recent census or other recent survey of the same population. Even in the absence of a recent census or survey, the number of boys and girls enumerated should be roughly the same. The number of children in each one-year age group should also be similar, with a slight decline in numbers with increasing age.

## CALCULATING THE INDICATORS

The basic tables necessary to calculate each indicator are listed in Appendix 4, and the programs to produce the estimates are given in Appendix 3. All the estimates can be produced with the programs found on the diskette supplied with this handbook. Otherwise, a standard package such as SPSS PC can be used.

You can obtain some of the indicators—for example, the per cent of households with iodized salt—directly from responses to variables on the questionnaire. For most indicators, however, you will have to examine the responses to a number of questions to determine the information needed in the indicator. For example, access to safe water is based on both source of water and distance to source. In these cases, a variable must be created indicating whether or not the household, mother *or* child fits the definition of the indicator.

☞ The EPI INFO program supplied with this handbook will have to be adapted to report on indicators that vary from country to country—for example, water and sanitation.

Most analysis packages, including SPSS and EPI INFO, will report a table of information on the variables created for the indicators. To report on the indicator, the appropriate number must be read off the table.

 **EXAMPLE:**

A table on low weight-for-age might be generated, as follows:

Value label	Value	Frequency	Per cent	Valid per cent	Cumulative per cent
≥ -2 SD	0	1,183	77.9	79.0	79.0
< -2 SD	1	314	20.7	20.9	100.0
	.	21	1.4	missing	
<b>Total =</b>		<b>1,518</b>	<b>100.0</b>	<b>100.0</b>	
<i>Valid cases: 1,497</i>				<i>Missing cases: 21</i>	

SD = standard deviation

The number to report for Indicator 11.1 in this case would be 20.9 per cent, the per cent of children whose weight-for-age was low, among those who had a valid weight-for-age in the data set. In the EPI INFO package included with this handbook, the indicators are printed directly using the REPORT function, so that only the necessary information is shown.

Anthropometric indices need to be calculated for each child, based on the measured weights and heights. EPI INFO uses a separate routine to compare a child's age and weight (or height) to an international standard, and then assigns a score for each child based on his or her deviation from the median value of the standard (a z-score). The calculations can be done at the time of data entry, as is done in the EPI INFO package supplied here, or can be done in a subsequent processing of the data using a special anthropometry package such as ANTHRO (formerly CASP) or EPINUT. The advantage of EPI INFO is that the indices can be viewed on the screen as soon as the height and weight data are input so that the data-entry staff can immediately detect indices which are implausibly high or low (usually due to keying errors when these data are entered). Indices below -6.0 z-scores or above 6.0 are generally considered to be implausible.

**EXAMPLE:**

The output generated by the REPORT program supplied with this handbook looks like this:

**Proportion of Children with Low Weight-for-Age Z-Scores**  
(Indicator 11.1,2)

<u>Total Children</u>	<u>WAZ &lt; -2</u>	<u>WAZ &lt; -3</u>
1,497	21%	4%

In some cases, the analysis package will not generate the indicator itself. For example, a survey will not give the absolute numbers in the population. To convert the per cent of the population with access to safe and convenient water (the number reported by the REPORT program) into absolute numbers (Indicator 13.1), it is necessary to multiply this percentage by the total population size.

### Estimating the Margin of Error for Each Indicator

Most analysis packages will estimate the standard error of percentages and means, assuming that the sample taken was a "simple random sample."<sup>1</sup> However, these standard errors are generally underestimated because they do not take into account the fact that the data come from cluster sample surveys. Analysis packages specifically designed to work with cluster sample surveys are needed to correctly estimate standard errors. EPI INFO (Version 6) has a module called CSAMPLE to compute correct standard errors for percentages and means, although this module is not integrated within the ANALYSIS module. Appendix 3 contains instructions for generating the standard errors and confidence intervals for each indicator using this module.

**The EPI INFO CSAMPLE program will automatically report the estimate along with its margin of error. See Appendix 3 for instructions on using this program.**

<sup>1</sup>The formula for calculating the *standard error* (s.e.) for a proportion, assuming a simple random sample, is:

$$s.e. = \sqrt{[p \times (1 - p) \div n]}$$

where  $p$  is the proportion of interest and  $n$  is the sample size in the group. The *margin of error* for this proportion is  $(2 \times s.e.)$ . The *confidence interval* for the proportion is calculated as  $p \pm$  the margin of error.

**EXAMPLE:**

The margin of error for the indicator and its 95 per cent confidence interval are shown. For instance, in analysing weight-for-age, the output was:

*Per cent with low weight-for-age is 21.0% with a confidence interval of (19.3, 22.7).*

### Calculating Weighting Factors for Non-Self-Weighting Samples

If separate sampling frames were used for different regions at the first stage of sampling, the national sample was not chosen with probability proportional to size (PPS). This may also happen if you stratified according to some other factor (e.g., urban/rural or slum/non-slum) and took different sampling fractions (proportions) in each stratum. These samples are not "self-weighting," and you must weight your sample when you report national estimates. That is, you must ensure that each separate subsample—for example, each separate region—contributes only what it would contribute if the survey sample at the national level had been chosen with PPS. The procedure for calculating weights is illustrated in Table 7.2 using a hypothetical example.

- ☞ Create a new variable to represent the weighting factor. This factor must be used to calculate *national* estimates from your separate subnational surveys. Each case will then be weighted by this variable.

For example, suppose the country has a population comprising 1,333,415 households and consists of seven regions, each of which you surveyed independently. The total sample of households was equal to 11,312.

To calculate the appropriate weighting factor for each survey sample, create a table such as Table 7.2.

- ☞ To calculate the actual coverage for the entire sample, construct a variable representing the weighting factor:

IF (REGION = 1) THEN WTFACTOR = 2.73

IF (REGION = 2) THEN WTFACTOR = 0.48

IF (REGION = 3) THEN WTFACTOR = 0.99

IF (REGION = 4) THEN WTFACTOR = 0.26

IF (REGION = 5) THEN WTFACTOR = 1.77

IF (REGION = 6) THEN WTFACTOR = 0.41

IF (REGION = 7) THEN WTFACTOR = 0.23

The analysis should then be weighted by the variable, WTFACTOR. The output will give national-level indicators properly weighted to allow for the non-self-weighting sample design.

☞ This calculation can also be done with a hand-calculator. Multiply each regional estimate by its weighting factor, add up the resulting products, and divide that sum by the sum of the weights.

📎 **EXAMPLE:**

(1) Region	(2) Estimate	(3) <u>Weighting</u> factor	(4) <u>Estimate x</u> WTFACTOR
1	0.85	1.7	1.445
2	0.75	0.2	0.15
3	0.70	<u>1.0</u>	<u>0.70</u>
Total		$\Sigma = 2.9$	2.295

National estimate weighted average: (sum of column 4 ÷ 2.9) = 0.79

**Table 7.2. Illustrative table for calculating weighting factors**

Stratum	(1) Population size (households )	(2) Stratum population as a proportion of national population	(3) Final sample size	(4) Stratum sample as proportion of total sample	(5) Weighting factor
		(col. 1 ÷ 1,333,415)	<i>n</i> (households)	(col. 3 ÷ 11,312)	(col. 2 ÷ col. 4)
1	550,088	.412	1,707	.1509	2.730
2	93,088	.070	1,664	.1471	.476
3	192,580	.144	1,648	.1457	.988
4	47,873	.036	1,560	.1379	.261
5	330,312	.248	1,587	.1403	1.768
6	77,124	.058	1,592	.1407	.412
7	42,350	.032	1,544	.1374	.233
Total	1,333,415	1.000	11,312	1.000	

---

### Sex- and Age-Specific Estimates

- ☞ To calculate each indicator by the sex of the child, you must select from the data file children of only one sex, and run the analysis for this group separately. Do the same for children of the other sex.

Follow the same procedure if you want to report indicators for narrow age groups, or for subnational groups such as regional, district or rural/urban groups. Select each group separately, and run the analysis program to get a report for one group. Repeat the process for each group for which you want separate indicators calculated. If you are not using the EPI INFO analysis program, you may find it easier to cross-tabulate the variable for sex of child, or other similar variables that define subnational groups (for example, regional, urban/rural).

- ☞ Make sure that you label your printouts clearly (e.g., "boys only") to avoid later confusion.

Now you are ready to write your preliminary report, and to display these indicators graphically.

## CHAPTER VIII

# REPORTING AND USING THE SURVEY FINDINGS

*This chapter is for programme decision-makers, survey coordinators and technical resource persons. It tells you how to:*

- ✓ Present the findings to decision-makers and the general public.
- ✓ Produce a preliminary report.
- ✓ Produce an executive summary.
- ✓ Produce a full report.
- ✓ Use your results. (Illustrations are provided.)

### REPORTING AT MID-DECADE

The report at mid-decade by the Secretary-General of the United Nations is important because it will summarise the way the international community has joined hands with governments, and they in turn with their constituents, to see that the promises made at the World Summit for Children engender action. The very act of measuring and reviewing progress toward the goals of each National Programme of Action is as important as achieving the goals themselves.

Each survey, from start to finish, is an opportunity for raising awareness. The measures of the mid-decade goal indicators should be used to reinforce programmes at all levels, to redirect or make corrective adjustments to programmes and National Plans of Action where necessary and to help policy makers take action that will build and sustain progress. That is why it is necessary to use the survey results to the fullest extent, building on the plans made in the early stages of survey planning.

### PRESENTING FINDINGS TO THE GENERAL PUBLIC

You and your team may be under intense pressure to release results quickly. Government officials, international organizations, statistical offices and even journalists will be asking, "What are the numbers?" These numbers, whether they are correct or not, will take on a life of their own once released, and may cause you a lot of trouble if you have not checked your arithmetic carefully.

During the course of the field work you should have used as many occasions as possible to report first impressions, to raise questions and to start corrective action at all levels. For example, you might ask the following questions: "Why isn't iodized salt found in the households?" "Why is

the vitamin A programme not reaching this area?" "Why are so many fewer girls attending school than reported by the school inspectors?"

Written summaries of overall results require more care. Once you have looked at your data critically, made the necessary calculations to produce your estimates and, if appropriate, plotted your results on graphs, you must produce reports that give "the numbers." These may take different forms, including policy briefs and summary sheets as well as reports.

Aim to get findings out quickly once you and your team have checked the results and tables. To speed results into action, plan on producing two reports: a preliminary report followed by a full technical report (which should include an executive summary).

Dialogue between health workers and community leaders can be encouraged by continued discussion around each report generated by this multiple-indicator survey. Even if subnational estimates are not reported, the national results can provide a focal point for discussing local or district-level problems and successes and for prioritizing future activities.


Raising public awareness of special problem areas (either specific indicators or areas at high risk) can lead to positive responses if the results are presented skillfully. And public understanding of problems can often be an important leverage tool, especially when programmes are community-based.

### **Producing a Preliminary Report**

The preliminary report is a brief and early version of the full technical report. It contains:

- the aims and objectives of the survey
- major results
- a discussion and preliminary recommendations
- the sample and methodology
- the field work: what questions were asked, how they were asked, and who asked them

Prepare the preliminary report first. Most sections of this brief report can be written as the survey operation progresses. Once the data are analysed, all that will remain will be to add the estimates of each indicator. Include only those major findings that are firm.

 It is important to have a section on the methodology even in the preliminary report. This section should describe the sampling scheme, the instruments, the field process, number of interviewers, losses and refusals and so forth. Position the methodology section in the report so that it does not divert the attention of policy makers away from the findings themselves.



Select the main findings and recommendations that you want everyone to remember. Use the preliminary report to circulate your main findings—some of which may require judgement and subsequent discussion—while there is still interest in the survey. If you have planned well and your survey organization runs smoothly, you should be able to prepare the preliminary report within one month to six weeks following the end of field work.

You may use the publication of the preliminary report as a vehicle to stimulate interest in the current health situation among government agencies, nongovernmental organizations (NGOs), other multilateral donors, the press and the public. You can also show where the biggest data gaps existed before the survey was conducted, and how these results have contributed to updating old information.

All major participants in the survey who are involved in the planning, implementation or analysis stages should be invited to participate in a meeting to present and discuss the main findings.

This publicity may result in additional interest in the survey results and lead to additional analysis of the data. Once the full technical report is complete, the survey data should be made available to qualified and interested parties for further use. This can contribute to the effective and widespread use of the survey results, and multiply the benefits of your effort and expenditure.

Other forms of survey reports may also be useful for reaching certain audiences: a short summary for those who were involved with the survey at some stage, such as interviewers or those who have given support to the survey; a report for community leaders about services and problems identified, including the most important conclusions of the survey analysis; and a press release summarising the most important conclusions of the survey. It is important to ensure that the press correctly reports the survey findings and does not distort them. Preparing your own press release can help ensure that results are accurately reported.

## **The Full Survey Report: Presenting Findings to Decision-Makers at Different Levels**

### ***The Executive Summary***

You and your national counterparts should agree upon the recommendations and actions to be taken upon publication of the final survey report. At this stage your findings are firm, and you and your team have had time to discuss and plan how to use them. This marks the start of advocacy efforts, programme consultations and strengthening and re-deploying resources.

Your full technical report should begin with an *executive summary* (see Box 8.1). Keep the executive summary short; it should not exceed a few pages, because policy makers will not read more than that.

The summary should include more descriptive and comparative information than technical information. It should report the highlights of the survey, but you must be sure of your findings since

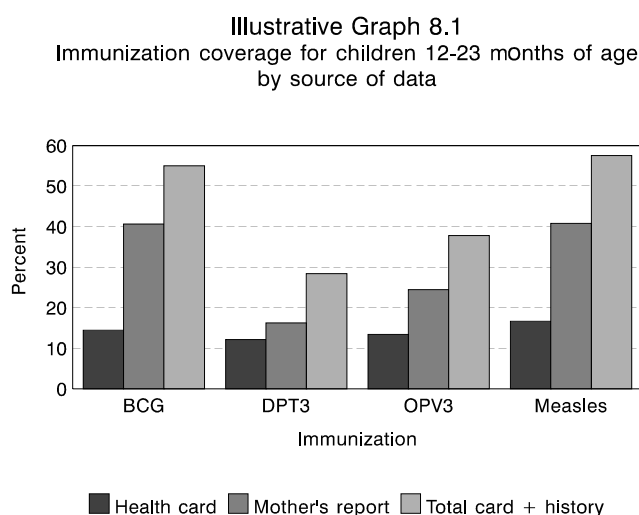
policy decisions will be made on the basis of this part of the report.

**EXAMPLE:**

Policy makers will want to know how the new estimates compare with past estimates. Has the country improved? How does it compare with other countries?

Remember that diagrams have more impact than long tables.

**EXAMPLE:**



The summary is also a tool for introducing the mid-decade goal indicators to a wide audience and can be used to emphasise the quality of previous information available about these indicators to policy makers. In many countries, some indicators have never been measured at the national level (salt iodization may be one example). Your presentation of results to policy makers should make this clear. The presentation should help to stimulate debate and lead to programmatic decisions where results indicate that action needs to be taken.

You may also be able to arrange special meetings on specific topics to which the survey results call attention.

**EXAMPLE:**

Legislation to promote salt iodization has been on the books for many years, and policy makers thought that public concern about goitre was unwarranted. The survey shows low levels of iodized salt in households—and that finding offers an opportunity to focus new attention on iodization needs and remedial action.

**Box 8.1**  
**ITEMS TO INCLUDE IN AN EXECUTIVE SUMMARY**

- ☞ Purpose of the survey
- ☞ Summary of progress in achieving goals, according to indicators measured by the survey
- ☞ Definitions of each indicator
- ☞ A brief description of the survey steps
- ☞ Country trends for goal indicators that have past data for comparison
- ☞ The main recommendations

***Producing the Technical Report***

The full technical report is essential if your survey results are to be viewed from an international perspective. This report enables all readers to judge the technical aspects of the survey operation, and to evaluate for themselves the quality of the data it produced and upon which your results are based. These surveys to monitor indicators of the Mid-Decade Goals must produce results that can withstand the heat of intense scrutiny and international comparison. For this reason, it is important to give all interested parties the information necessary to evaluate your results (see Box 8.2). That is the purpose of a full survey report.

The technical report is also the document governments will subsequently share for international reporting. Your role as an advisor is important for encouraging governments to share their results. An emphasis on process rather than on achievements should balance pride in accomplishment with an interest in turning problems into reasons for creating alliances for action. These alliances can be forged at many levels, including regional and international ones, when reports are shared openly. As soon as possible, copies of the report should be transmitted to all interested parties, including your UNICEF Regional Office and UNICEF Headquarters.

Be sure to include a discussion of the quality of the data the survey produced. Report the number of households sampled that did not complete the questionnaire, indicating the reasons for nonresponse. Indicate which items in the questionnaire were particularly difficult for respondents to answer, indicating those with high proportions of respondents answering "don't know" or items with a high proportion of responses coded as "missing." Also, describe your survey sample of households, women and children according to the sociodemographic characteristics you collected. For example, describe the per cent urban and rural, the average number of persons per household, the sex and age distribution of children and any other characteristics of the sample you think are important. Describe the details of the sample design in an appendix to the main report.

**Box 8.2**  
**THE COMPONENTS OF A FULL SURVEY REPORT**

- Executive summary.
- Aims and objectives of the survey.
- Sample design and size, what was used as a sampling frame, and the method used to choose households in each community.
- The questionnaire.
- Description of the field work.
- A description of when and how the field work was done and a list of the supervisors and interviewers, including a copy of the questionnaire, along with interviewers' and supervisors' instructions.
- Data entry procedures:
  - Was data cleaning and editing necessary? What was done?
  - How much time and how many personnel were needed?
- Evaluation of data quality, including a description of problems with missing data for any particular questions and the nonresponse rate.
- Results and interpretation of your estimates.
  - Where possible, compare your results with those obtained from other sources and with neighboring countries.
- Conclusions and actions planned, recommendations for future surveys or studies, acknowledgements for help and support received from individuals and organizations.

***Report the Indicator Estimates***

Describe briefly each method of calculation and the data used in the calculations. Report the values of each indicator. Explain and report the margin of error around your estimates. Caution your readers that all the estimates obtained fall within these confidence limits.

Use tables, graphs and charts to display your results.

 **EXAMPLE:**

Include a table giving low-weight-for-age data, as in the following illustrative table:

**Illustrative Table 8.1 Per cent of under-fives with low weight-for-age (and margins of error), urban and rural areas**

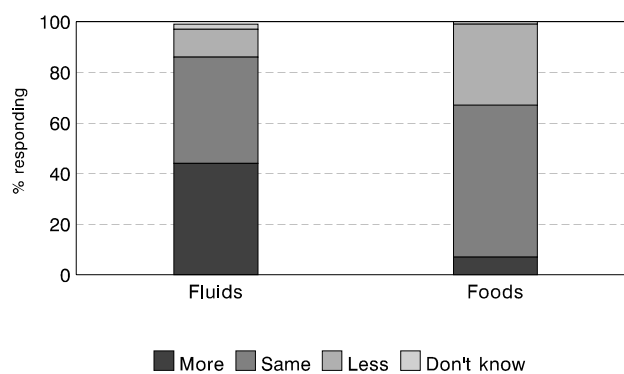
Residence	Low weight-for-age (%)*	Number of children
Urban	10.7 (9.5, 12.5)	1,273
Rural	15.9 (13.0, 16.7)	1,989

**Note:** Children are classified as low weight-for-age if their z-score was two or more standard deviations below the median international standard.

**EXAMPLE:**

Present a graph showing the changes in the amount of fluids and food that mothers give to children who have diarrhoea.

**Illustrative Graph 8.2**  
Changes in the amount of fluids and foods given to children with diarrhoea



Note: Diarrhoea during two weeks preceding survey

Report estimates for subpopulations if your sample size allows these calculations.

**EXAMPLE:**

Report estimates for males and females separately, for urban and rural areas separately (if you have distinguished rural from urban clusters) and for regions.

Compare your results with any earlier estimates and reports which are available to you. Read reports from other surveys before trying to interpret your results and point out any caveats about the comparisons being made.

**EXAMPLE:**

Present a table giving information about diarrhoea prevalence.

**Illustrative Table 2. Incidence of diarrhoea by child's sex and residence, and per cent of children under five years of age who had diarrhoea in the two weeks preceding the survey**

Variable	Per cent with diarrhoea	Number of children
Sex		
Male	27.5%	2,582
Female	28.7%	2,580
Residence		
Urban	28.1%	2,535
Rural	28.1%	2,626

Finally, be sure to circulate an early draft of the report to all collaborators and others who may be able to provide valuable comments before it is finally published.

## FROM DATA TO INFORMATION

### Using the Bangladesh Survey Data

With the Bangladesh Bureau of Statistics (BBS), the Institute of Statistical Research, Bangladesh, and the UN Statistical Office, UNICEF's Bangladesh office helped develop and implement a Multiple-Indicator Survey that helps to monitor many of the Mid-Decade and Year 2000 Goals. Seen as a valuable complement to other data collection activities and a tool for validating routine reports from the health services, the methodology has been institutionalised for continued use by the BBS. Two surveys, in 1993 and 1994, have already been done and will be repeated annually by the BBS.

From the very beginning of the survey planning process, special emphasis was given to the use of the survey results. This meant that at the very start, selection of the questions to be asked in the survey was tied very precisely to the uses that would be made of the answers.

In Bangladesh the results of the Multiple-Indicator Surveys have served a myriad of purposes, making full use of the time and effort which went into planning and carrying out these surveys. In the following sections, some of the key areas in which the data have been put to action

are described. Goal-monitoring has become a self-sustaining process in Bangladesh, with the government initiating and directing these simple surveys. Not only have UNICEF programmes benefited from the survey data, but so have many other programmes supported by other agencies. This diverse and widespread use has been possible because the Bangladesh office made certain to work with national partners to build capacity and to see that results were given wide exposure. Policy reports made the information accessible to other organizations in readable and usable form. It is clear that the methodology and the results of the survey are fully integrated into the social development schema of Bangladesh.

---

---

**The surveys promote improved national capacity for assessing the situation of children and tracking progress toward national goals for children.**

---

---

### ***The Final Report: Monitoring the Goals***

The survey results are used to report on the progress made toward achieving goals in the country. They are disseminated in the book *Progothir Pathey*, meaning "the path of progress," which is published annually by the BBS and is widely distributed to all ministries, district officials, NGOs, donors and academic institutions. It contains national trend information on the mortality, fertility and literacy indicators; maps showing the spatial coverage variations of selected indicators; summary tables of data disaggregated by gender, area of residence (urban/rural, slum/non-slum) and ethnicity; and district-specific tables reflecting the achievement of various indicators by gender. *Progothir Pathey* has already become a reference publication for researchers, social scientists and development workers in Bangladesh. The results of the survey are also being used to validate the completeness of the routine reporting systems of the government.

### ***Other Reporting Formats: Advocacy and Policy Dialogue***

The results of the survey provide valuable information for initiating advocacy and policy dialogue, not only for UNICEF, but for other UN and donor agencies and NGOs. The Center for Policy Dialogue, an NGO, is currently carrying out an in-depth analysis of survey data to initiate a national policy dialogue of some of the key disparity issues. The employment of maps to show the use of iodized salt in the country has been critical in shaping government policy for importing iodized salt and enforcing the universal salt iodization legislation in the country. UNICEF is also using these results to publish their annual statistical card series—a set of pocket cards containing key data, their implications and desired actions. These statistical cards are very popular amongst political leaders, administrators and the donor community.

### ***Using the Survey Results for Expanded Action***

The results of the survey are compared with the Mid- and Year 2000 Goals to calculate a "challenge factor," which is the gap between the goal and the existing level of coverage, expressed as a per cent. The challenge factor is then used to classify districts by their level of performance and therefore by the level of attention and inputs necessary to bring them up to the level of other districts. The annual average rate of change (AARC) is then used to set district-specific targets for various indicators. The survey data are further analysed to calculate the design effect for each indicator in each district. This design effect indicator is used to identify areas with the greatest intra-district variations in levels of coverage. Based on this analysis, key district-specific strategies for action are identified. All this information is presented in a situation analysis published and prefaced by the Planning Commission in Bangladesh. *KAMONA* ("desire"), as it is called, is distributed to all ministries, district-level staff, NGOs and donor agencies. A Bangla version of *KAMONA* is distributed as a part of the State of the World's Children report kit. It provides a local-level agenda for action for social development workers in Bangladesh.

### ***Subnational Planning: Looking Beyond National Averages***

Data disaggregated by gender and area of residence have brought the need for subnational planning to the forefront. With the help of the survey results, it has been possible to identify the districts which are lagging behind in various sectors. This information has resulted in the development of more efficient and need-based plans of action in particular areas for the achievement of the Mid-Decade and Year 2000 Goals.

**The survey should encourage policy makers to look at the information behind the national averages.**

### ***Interagency Cooperation***

The survey results have provided another opportunity for interagency cooperation. The availability of survey data and its use for synthetic estimation of mortality and fertility at the district level now allows UNICEF and UNDP to work together to develop district-specific human development indices (HDIs) for inclusion in the *Bangladesh Human Development Report, 1995*. The Asia Development Bank (ADB) and the World Bank have also asked UNICEF for support in organizing special surveys, using the Multiple-Indicator Survey methodology, for collecting subnational data necessary for project design. This cooperation opens new possibilities

**The surveys provide special opportunities for promoting interagency cooperation between international organizations and the host government.**



for UNICEF's participation in the design of large projects and as well as in policy dialogue.

### ***UNICEF Country Programme of Cooperation: Increased Efficiency and Effectiveness***

The UNICEF Country Programme of Cooperation 1996–2000 uses the Convention on the Rights of the Child as its framework. With most of the Mid-Decade Goals and some Year 2000 Goals already achieved in Bangladesh, the new programme places emphasis on equity and reduction of disparities. The survey data have successfully identified interdistrict, interdivision, gender and ethnic variations and form the basis for the design and development of programmes for 1996–2000. The phasing plan of the Accelerated District Approach (ADA) of the new proposed country programme uses challenge factors to identify districts requiring the most attention and input. Thus, districts with the highest challenge factor—the weakest districts—are given more time, input and attention to help them to catch up with the other districts in levels of coverage and performance.

---

---

**Survey results can help UNICEF make sure that its own programming efforts and its assistance are more efficient and effective.**

---

---

### ***Staff Assignment and Budget Exercise***

The staff assignment and relocation for the 1996/97 budget exercise is also based on the results of the Multiple-Indicator Survey. For total office mobilisation to make an impact in achieving the goals, senior and more experienced staff have been assigned to districts with high challenge factors. Because it recognises the complexity of working in a high-challenge factor district, UNICEF Bangladesh has limited the responsibilities of the staff in those areas to facilitating the achievement of the goals in only one district, while staff assigned to low-challenge factor districts might be covering more than one district. This approach has allowed UNICEF Bangladesh to more rationally reallocate its staff (and even retrench some posts) without affecting overall office performance and efficiency.

### **Reviews of Progress by Heads of State, Mayors and the Public**

In country after country attention at the highest level of policy-making to the needs of children and women has turned goal-setting into visible improvements. There is a demand for current, reliable information to assess problems and progress. The time-bound, measureable nature of the Mid-Decade and Year 2000 Goals has set a premium on responding to the demand to know.

In the past, surveys or other monitoring systems all too often meant a lengthy, costly process where the results came long after policy makers had posed the questions. The desire to know, and

the potential for action, had then passed. Final reports resembled academic treatises, discouraging even those dedicated to learning from the results.

Today, high-level policy concern, skilled summaries by planners and the media as well as action by programme managers are making it possible for the public to demand and learn the progress being made toward goals.

 **EXAMPLES:**

The president of Mexico has presided over six public meetings to evaluate progress on the National Programme of Action.

Mayors in the Philippines in the 60-member League of Cities initiated annual reports of progress to track the impact of City Plans of Action for children to achieve the goals.<sup>1</sup>



This handbook provides you with the means to work with others to produce current assessments of progress in a focused, rapid, cost-effective fashion. We hope you will find it a useful resource. Each survey provides an opportunity to raise awareness, foster dialogue between health workers, community leaders and public officials, and to take corrective action. Measuring and presenting current status on the indicators is a most powerful driving force toward informed action for children.

---

<sup>1</sup>For a summary of national and subnational action and status in the Philippines, see *Progress of the Philippines: 1995 Mid-Decade Goals* (Manila: Government of the Republic of the Philippines and UNICEF, 1994).

**Box 8.3****Checklist for Reporting Your Results**

- **Produce a preliminary report.** Include:
  - ✓ aims and objectives of the survey
  - ✓ major results
  - ✓ discussion and preliminary recommendations
  - ✓ sample and methodology
  - ✓ field work: questions asked, how they were asked, and who asked them
- **Produce an executive summary.** Remember to:
  - ✓ explain why mid-decade goal indicators are measured
  - ✓ be sure of your findings
  - ✓ include more descriptive and comparative information than technical information
  - ✓ summarise progress in achieving goals, according to indicators measured by the survey
  - ✓ give definitions of each indicator
  - ✓ describe the survey steps
  - ✓ provide the country trends for goal indicators that have past data available for comparison
  - ✓ present the main recommendations
- **Produce a full technical report.** Include:
  - ✓ an executive summary
  - ✓ aims and objectives of the survey
  - ✓ details of the sampling techniques used
  - ✓ a copy of the questionnaire used
  - ✓ details of the field work
  - ✓ data-entry procedures
  - ✓ an evaluation of data quality
  - ✓ the results
  - ✓ an interpretation of the results, comparing them with results from other sources, and with data from neighboring countries
  - ✓ conclusions, recommendations and acknowledgements
- **Make the reports of survey results fully accessible to all organizations which might make use of them.** Ensure the widest possible dissemination of results by calling discussion meetings with donors, ministry officials and community leaders, and by ensuring press coverage of these meetings.

# APPENDIX 1

## **SURVEY MANAGER:**

*Give a copy of this appendix to every interviewer. Translate this appendix into the interviewers' local language if necessary. Ensure that the information in this appendix is appropriate to your survey. Some countries may need to give slightly different instructions.*



## Instructions for Interviewers

### CONTENTS

Introduction .....	A1.4
How to Handle an Interview .....	A1.4
General Points .....	A1.5
How to Fill in the Questionnaire .....	A1.6
Household Module .....	A1.6
Water and Sanitation Module .....	A1.9
Salt Iodization Module .....	A1.9
Education Module .....	A1.9
Modules for Mothers .....	A1.10
Tetanus Toxoid (TT) Module .....	A1.10
Care of Acute Respiratory Infection Module (optional) .....	A1.11
Modules for Children under Five Years of Age .....	A1.11
Diarrhoea Module .....	A1.11
Vitamin A Module .....	A1.12
Breastfeeding Module (optional) .....	A1.13
Immunization Module .....	A1.14
Alternative Immunization Module .....	A1.14
Anthropometry Module .....	A1.14
Mortality Module .....	A1.16
Filling in the Cluster Control Sheet .....	A1.17

## INTRODUCTION

The principal aim of this survey is to monitor the progress of health programmes. This will be done by asking mothers or other caretakers of young children about the health of the children they care for. Detailed information will also be collected on the household and on the mothers' own immunizations and knowledge of health matters.

Your supervisor will have given you a list of households to visit, or will tell you how to find the households. You must visit all these households. In each household you should interview all the mothers or principal adult caretakers of children under age 15 years. If you visit a household with no children under age 15, you must still ask questions about the household. You may ask these questions of any adult who is present.

If no one is at home when you go to interview the household, ask the neighbours whether the house is inhabited. If it is occupied, ask the neighbours when the household members will return. Arrange with your supervisor to go back to the dwelling when it will be occupied, or at the end of the day. Note those plans on your cluster control form and note the time you are to call back on the household listing page of the questionnaire. Do not substitute another household.

If no adult is at home, arrange to come back at another time. Do not interview a temporary caretaker of the children, such as a baby-sitter.

If nobody is at home again when you make the return visit, then mark this household on your form as "missing." Your supervisor may instruct you to revisit these households, or to visit another household which has been chosen at random.

In areas where separate households are difficult to identify, such as multi-household dwellings, treat the entire dwelling as one household, and interview all the women and children within that dwelling.

### **EXAMPLE:**

If the house or dwelling chosen is an apartment house and contains several households, all the households should be interviewed.

Ask your supervisor if you are in doubt about what to do when you cannot locate a household, or you cannot complete an interview. Always keep a record on the cluster control sheet of the households you visited where nobody was at home.

## HOW TO HANDLE AN INTERVIEW

- Conduct yourself in a relaxed informal way, but be thorough. Use the questionnaire carefully.
- Ensure that you understand the exact purpose of each question. This will help you to know if the responses you are receiving are adequate.

- Ask the questions exactly as they are written. Even small changes in wording can alter the meaning of the question.
- Ask the questions in the same order as they are given on the questionnaire.
- Ask all the questions, even if the respondent answers two questions at once. You can explain that you must ask each question individually, or say "Just so that I am sure . . ." or "Just to refresh my memory . . .," and then ask the question.
- Help your respondents to feel comfortable, but make sure you do not suggest answers to your questions. For example, do not "help" a woman to remember the age of her child.
- Do not leave a question blank unless you have been instructed to skip the question. Questions left blank are difficult to deal with later. In the office it may look as though you forgot to ask the question. Always write in 0 when a zero answer is given. Write in the code number (usually 9 or 99) for responses such as "not applicable" or "don't know," when necessary to indicate that you asked the question but were not given a proper answer.
- Record answers immediately. Write down any pertinent remarks made by other people present, and mention who the other people are. Check the whole questionnaire before you leave the household to be sure it is completed correctly.
- Thank the respondent for her cooperation. Remember the survey schedule and do not stay and talk for too long.

## **GENERAL POINTS**

### ***Dress neatly.***

Your appearance gives the respondent's first impression of you. The way you dress will affect whether your interviewing is successful or not. Dress neatly and simply.

### ***Gain rapport with the respondent.***

Try not to arrive at the respondent's house at an inconvenient time of day, such as a meal time. Try to arrive when the respondent will not be too busy to answer questions.

Introduce yourself by name and show your identification.

Explain the study and why you want to interview the women in the household, exactly as your introduction tells you to.

Be prepared to explain what is meant by confidentiality and to convince respondents to participate if they are reluctant.

If the respondent refuses to be interviewed, note the reasons on the questionnaire, if possible.

Remain calm and polite at all times.



***Probe for adequate responses.***

Pause and wait if the respondent is trying to remember difficult items.

Ask the respondent to clarify her answer if necessary. You may be misunderstanding the respondent.

Check for consistency between answers a mother gives. For example, if she tells you that her child received only breast milk yesterday, she should not answer "yes" to the question "Since yesterday, has she been given anything to drink from a bottle with a nipple or teat?"

**HOW TO FILL IN THE QUESTIONNAIRE**

The following instructions refer to the questionnaires that appear in the Model Questionnaire Modules at the end of chapter 3. Translate these instructions into the language of the interviewer.

**HOUSEHOLD MODULE**

The Household Module, which appears on page Q1 of the Model Questionnaire (at the end of chapter 3 in this manual), has two parts: (1) a Household Information Panel (replicated below) for recording household-specific information and (2) a Mother and Child Listing Form for listing mothers and their children or caretakers and the children for whom they are responsible. Fill out the Household Information Panel first.

**Household Information Panel**

- (a) Enter the cluster number as instructed by your supervisor.
- (b) Enter the household number as instructed by your supervisor.
- (c) Enter the date of the interview: day/month/year.
- (d) Enter your own name or identifying number.
- (e) Enter the name of the head of household (HH).
- (f) If household is unoccupied, enter time to re-visit, as instructed by your supervisor.
- (g) Enter the number of persons who usually live in the household—that is, the total number of persons who usually live and eat together.
- (h) Enter the correct code for the material of the dwelling floor. You will be able to observe the correct answer in most cases, but if in doubt, ask. If there is more than one kind of material making up the floor, record the main flooring material.
- (i) Ask and enter the number of rooms in the dwelling.
- (j) The data entry clerk should fill in this cell at the time of data entry.
- (k) Enter codes for region and area of residence—that is, urban (1) or rural (2)—here, as instructed by your supervisor.
- (l) Fill this section in once you have completed all the modules. If all household modules have been completed, all mothers of under-fives have been interviewed, and questionnaires for each child under age 5, and the education module for all over-fives, has been completed, note that the interview is complete here.
- (m) If you are unable to complete the interview, enter the reason.

**Household Information Panel**

(a) Cluster number:	(b) Household number:	(c) Date of interview (day/month/year): __ / __ / __
(d) Interviewer no.:	(e) Name of head of household:	(f) Call-back needed? Yes / No Time: __ A.M./P.M.
(g) No. persons in HH usually resident:	(h) Material of dwelling floor: 1 wood/tile 2 planks/concrete 3 dirt/straw 4 other	(i) Number of rooms in dwelling:
(j) Data entry clerk no.:	(l) All forms completed? 1 Yes 0 No	(k) Region 1 2 3 4
	(m) If not, why not? 1 Refusal 2 Not at home 3 HH not found/destroyed 4 other	(k) Urban 1 Rural 2

**Mother and Child Listing Form**

Next, fill in the Mother and Child Listing Form (bottom of page Q1 in the Model Questionnaire). The form provides questions to obtain the names and ages of mothers in each household, and to obtain the name, age, birth date and sex of each child who is usually resident in the household. It will also provide a record of your visits and the forms you complete for each household.

Ask to speak to the mothers or others who care for the children in the household. If a mother is not at home, direct the questions to the person caring for her children. In the last column on the listing form, indicate if the respondent is the mother (3) or caretaker (4). If no mother or adult caretaker is at home, tell your supervisor, and decide what to do about re-visiting the household.

**Line no.:** Fill in the line number 1-0 (the first line in the Child Listing Form) to refer to the first mother interviewed in the household. Fill in line numbers for each child of that mother, starting with 1-1.

**Question 1. Name:** List the name of the first mother or caretaker in Line 1-0. List the names of each of the children she cares for, starting with the youngest in Line 1-1.

**Question 2. Male/Female:** Circle 1 if the person listed on that line number (caretaker or the child) is male, 2 if female.

**Question 3. Date of birth:** Ask for the child's date of birth. Use a 0 in any spaces if the month or day does not contain 2 digits—for example, the month of March is coded as "03." For children under age 5, it is important to record the child's month and year of birth accurately, especially if the Anthropometry Module is included in the survey. For children age 5 and over, the age in completed years is all you need to record. You do not need to enter the mother/caretaker's date of birth.

If the mother can give a year of birth, but not the month, probe to try to estimate the month. You will have

to convert month to a number, as you have been instructed. If the mother cannot give a year and month of birth, ask to see the child's birth certificate or vaccination card, and write the date given on the card. Check this date with the mother to be sure she thinks this date is accurate. If you cannot estimate the child's month of birth from this information, you may need to find out in which season he/she was born. You may need to go on to question 4 in order to estimate a year and, if possible, a month of birth, and then to enter it here. If it is still not possible to estimate the child's month of birth, write "99" in the space for MONTH. If you cannot ascertain the DAY of birth, write "99" in the space provided for DAY. You must record a YEAR of birth. Do not leave these columns blank.

If the mother cannot give a year of birth, you may need to use a calendar of important events to estimate the child's year of birth. If there is no other way to estimate the year of birth, relate the child's birth to another child in the household, for whom the date is known. You will need to use your judgement to estimate a year of birth for each child.

**Question 4. Age:** Ask for the age of the mother/caretaker and enter in the row with his/her name.

Ask for the age of each child. Enter the response in completed years. If a child has not yet reached his/her first birthday, write "00"—for example, a child who is 9 months old is coded as "00."

You may be instructed to compute the child's age in months or years from the date given (or the birth certificate or immunization card), using the conversion chart you were given in training. (See example of conversion chart found at the end of this appendix.) You will be given instructions about how to do this during your training. You will then be asked to check that the age you compute from the date given corresponds to the mother's report. If the two ages match, enter the child's age in completed years here. If they do not match, you will have to probe further to estimate the child's correct age.

Remember, ages must be at last birthday; that is, the child (or mother/caretaker) must have already passed or celebrated his/her birthday.

If the mother does not know the date of birth or current age of her child, try asking "How many years ago was [name] born?" You can help her by relating the child's age to that of other children or to some important event or to the season of birth, by asking, for example, "How many wet seasons ago was [name] born?"

**You must record an age for each child. Do not leave this column blank for any child.**

Stop listing when the woman reaches a child who is over the age of 15 years.

**Question 5. Category of respondent:** Circle "1" if the child is 5 years or older; circle "2" if the child is under age 5. Circle 3 if the respondent is the mother and 4 if she/he is a caretaker. You will use this information to decide which modules to administer to children in the household.

When you have finished with the first mother, go on to list the next mother/caretaker in line no. 2-0. Then, list her children, starting with the youngest in line no. 2-1.

Then ask "Are there any other children who live here, even if they are not at home now?" Continue listing children in the household, adding a continuation page if necessary. Tick the space provided on the first listing sheet if a continuation page is attached.

### **WATER AND SANITATION MODULE**

**Question 1.** The purpose of this question is to find out how safe the household water used for drinking is. If several sources are mentioned, probe to determine the most usual source. Enter the code for the most usual source.

**Question 2.** This question is used to find out how convenient the source of drinking water is to the dwelling. Enter the code for the answer.

**Question 3.** This question also aims to find out how convenient the water source is. Fill in the estimated time (in minutes, converting from hours, if necessary) it takes by the usual mode of transport to get to the water source, wait to get water, and get back to the dwelling. If the water source is on the dwelling premises, record "on premises" by entering 888. Enter "999" for the response "don't know."

**Question 4.** The purpose of this question is to obtain a measure of the cleanliness of the sanitary facility used by the household members. If the respondent answers that the household members use the bush or field, enter "9"—No facilities. If "No facilities" is the answer, skip question 5 and go to next module.

**Question 5.** The purpose of this question is to assess the convenience of the sanitary facilities used by the household. Enter the code number for the response given.

### **SALT IODIZATION MODULE**

**Question 1.** This item is used to record the type of salt used to prepare the family's main meal the day before the questionnaire is administered, and the outcome of the test for salt iodization. Follow the instructions given in training to test the salt, and record "1" (iodized) if the test is positive (>20 ppm); "2" if the test is negative. Circle "3" if no salt is found in the home and "9" if the salt was present, but not tested for any reason.

**Question 2.** Record the type of salt found in the home.

### **EDUCATION MODULE**

The questions about school attendance are asked to determine attendance at primary school. These questions are asked about all children age 5 or over (or over the usual age at school entry).

List names and line numbers of all children 5 years and older at the top of this page. See Column 6 on the Child Listing Form on the first page of the questionnaires, and list all children for whom a "1" is circled on this page (that is, "Yes" to Question 5, "Age 5 or over?").

If the survey is collecting data between the end of one school year and the start of the next, Question 2 should apply to the year just finished.

**Question 1.** For all children age 5 or over (or usual school entry age), ask "Has [name] ever attended school?" Circle 1 if the answer is "Yes" and 0 if the answer is "No." If the respondent does not know the answer, circle "9." If 0 or 9 are circled, skip to the next child. If no other children over age 5 live in the household, skip to the next module.

**Question 2.** If the answer to Question 1 is "Yes," ask: "Is he/she currently at school this year?" and circle the code for the response given. If the answer is 0 or 9, skip to question 4.

**Question 3.** Fill in the child's current grade and code number for the level of school attended.

**Question 4.** If the child attended school last year, circle "1" and go on to question 5. If not, circle 0. If the respondent does not know, circle 9, "Don't know." Go on to next child if 0 or 9 is circled. If no other children over 5 reside in the household, go on to next module.

**Question 5.** If the child attended school last year, fill in the child's grade and code for the level of school attended last year.

## MODULES FOR MOTHERS

The Modules for Mothers (beginning on page Q4 in the Model Questionnaire) consist of the Tetanus Toxoid Module and the Care of Acute Respiratory Infection (CARI) Module for mothers of children under 5 years old.

### TETANUS TOXOID (TT) MODULE

Fill in column headings with line number and name of each mother of a child under 5 listed in the Household Module. Do not administer this module to caretakers—only to the biological mother of the child.

These questions are directed to mothers of under-fives, and aim to obtain information about whether a child was protected against tetanus at his or her birth.

**Question 1.** First, ask whether the mother has a vaccination card or other documentary evidence of vaccination. Ask to see this evidence and record that you have seen it.

If no card is available, you must try to find out how long ago the last TT dose was received, and the total number of TT doses the mother has received in her lifetime. Use the probing questions, and fill in her answers in the boxes provided.

**Question 2.** Ask if she received any tetanus toxoid injection during her last pregnancy. Record 1 for a "Yes" response and 0 for a "No" response to question 2.

**Question 3.** Enter the number of doses the mother reports in the space provided. If the mother reports two doses during the last pregnancy, stop here.

**Questions 4 and 5.** If she reports fewer than two doses, continue to probe for earlier doses of TT. Ask about doses received during or before the next-to-last pregnancy or between pregnancies. Enter her responses in the spaces provided, as in questions 2 and 3 above.

**Question 6.** If the mother received no dose during her last pregnancy, but had received earlier doses of TT, ask her to estimate how long ago the last dose was received, and enter her response (number of years ago or month and year of last dose) in the box for question 6.

**Question 7.** Then total up the number of doses the mother reports (add doses in boxes 3 and 5) and enter the total in the box next to question 7 (total doses received in lifetime).

#### **CARE OF ACUTE RESPIRATORY INFECTION MODULE (OPTIONAL)**

This question aims to find out if the mother knows when to take a child with a cough or cold to a medical provider. Fill in column headings with the line number and name of each mother of a child under 5 listed in the Household Module. Ask the question once for each mother/caretaker of children under 5 years in the household.

Do NOT suggest answers to the mother. DO NOT PROMPT. If the mother gives a response that does not appear on the list, write it in the space provided (1G). When all mothers of young children (under-fives) in the household have been questioned, go on to next module.

#### **MODULES FOR CHILDREN UNDER FIVE YEARS OF AGE**

Go to the Mother and Child Listing Form on page Q1, and, starting with the youngest child, fill out a separate child health questionnaire for each child under age 5 years, recording the name and line number of each child. Be sure to fill in the child's identifying information at the top of each page of the Modules for Children Under Five Years of Age.

#### **DIARRHOEA MODULE**

These questions aim to find out if the child has recently had diarrhoea and, if so, what treatments, drinks and foods the child took during the episode. A diarrhoea episode is defined as three or more loose or watery stools per day, and/or blood in the stool on any day, or as defined by the mother.

**Question 1.** Record the mother's answer. If a respondent is not sure what is meant by diarrhoea, tell her it

means "more than three loose or watery stools per day or blood in the stool." Make sure the respondent understands what is meant by "in the last two weeks."

If the child has had diarrhoea in the last two weeks, enter 1 ("Yes") and go on to question 2. If the child has not had a recent episode of diarrhoea (within the last two weeks) enter 0 ("No") or 9 ("Don't know"), skip the rest of the questions in this module and go on to the next module.

**Question 2.** We want to know what the child took to drink during the last episode of diarrhoea. Ask each question separately: Was [name] given breast milk? Was he/she given cereal-based gruels? and so forth. Circle the correct code for each answer: 1 = Yes, 2 = No, 9 = Don't Know. You will be given a list of "locally defined acceptable home fluids," like sugar-salt solution, yogurt drink, and so on. If the respondent answers "Water," ask if this was the only drink offered, or if it was accompanied by other feeding. Ask about other fluids considered "unacceptable" by the local Programme (you will be given a list of these). If the child took nothing to drink (2I), skip to question 4.

**Question 3.** If dehydrated, a child may take more fluids than usual. We want to know if the pattern of fluid consumption changed during the episode. Ask the question just as it is worded here. Read out the entire question, and enter the code for the mother's response in the box at the right.

**Question 4.** A child may change the amount usually eaten during a diarrhoea episode. Ask the question just as it is worded on the questionnaire. Make sure that the mother understands that this includes breast milk, if the child is still being breastfed. If the mother's answer is "less than usual," probe to try to find out if she thinks the amount was "somewhat less" or "much less" than usual. Record the code for her response in the box.

## VITAMIN A MODULE

Some questions are asked about the child's household, and are asked only once of each mother. Only the relevant sets of questions (Module A, B and/or C) will appear in the questionnaire.

**Question A1.** This question asks if the child has received a supplement. Show the sample capsule or dispenser you were given to help the mother remember. (You may be instructed to show several different capsules, asking the mother to identify the correct one.) Record the code for her answer. If the child has not received a vitamin A supplement, or the mother/caretaker does not know if he or she has received one, skip to next module.

**Question A2.** If the answer to question A1 is "yes," record the number of months ago the last dose was given to the child.

**Question B1.** This question asks if the fortified food product is used in the household. Show the sample packet (if given). If the answer is "yes," ask to see the household's packet, recording "2" if seen. If the answer is "no," record 0 and go on to next module.

**Question B2.** This question asks if the child has received the fortified food *during the last week*. Substitute the day of the week; that is, if the day you are interviewing is Thursday, ask: "Since last *Thursday*, has *Fatima* been given *fortified sugar* to eat?" Enter the code for the answer. Prompt if necessary.

**Question C1.** This question asks if the mother has heard the educational programme messages about vitamin A-rich foods. Record her answer in the space provided. If the answer is "no" or "don't know," SKIP TO NEXT MODULE.

**Question C2.** If the answer to question C1 is "Yes," this question asks the mother to list the foods she has heard of. Enter "1" if she mentions any of the correct [*country/region/season-specific*] vitamin A-rich foods promoted by the programme listed on the questionnaire in question C2a and Question C2b. Enter "0" if she does not mention the food listed. If she gives any other response, enter 1 for question C2c. A mother's answer is coded as correct if at least one "1" response is entered.

**Question C3.** This question asks if the child ate the specified foods during the past week.

Substitute the day of the week; that is, if the day you are interviewing is Thursday, ask: "Since last *Thursday*, has *Fatima* been given [*vitamin A-rich food*] to eat?" For each eligible child, record the mother's answer.

#### **BREASTFEEDING MODULE (OPTIONAL)**

**Question 1.** This question asks if the child has ever been breastfed. It includes any breastfeeding experience of the child—not necessarily by the mother. If the answer is "Yes," enter "1" and go on Question 2. If the answer is "No" or "Don't Know," enter the correct number and Go on to the next module (or optional question 4).

**Question 2.** Enter code for mother's response; if the child is no longer breastfed, go on to the next module or to optional Question 4 and/or Question 5.

The next questions ask about what the child was fed in the preceding 24 hours.

**Question 3.** The purpose of this question is to determine what other liquids or foods the child was given in addition to breast milk. Ask each possible answer in turn—for example, "Did [name] receive vitamin or mineral supplements or medicine? Did he/she receive plain water?" etc. Circle the appropriate code for each answer: 1 = yes, 2 = no, 9 = don't know. If a mother mentions another liquid or food which does not fit into one of these categories, write it in the space provided in 3F and circle "1."

**Question 4 (optional).** This questions asks if, in the last 24 hours, the child was bottle-fed.

**Question 5 (optional).** This question asks when breastfeeding for this child ended. It is used only in countries with very short breastfeeding durations. Record the age in months when breastfeeding stopped. Record "99" if the mother does not know or cannot provide an answer.



### IMMUNIZATION MODULE

This module is used to obtain the numbers of children between 12 and 23 months of age who received BCG, DPT3, OPV3 and measles immunization before age 1. Here, the information is obtained for all under-fives in the household. First ask if the child has a vaccination card, and record the correct code in the box at the right. If yes, ask to see the card. Copy dates of all relevant vaccinations to the space provided. If not date for vaccination is recorded, or if no card is available, use the probing questions to ask the mother whether the child has received each vaccine. Enter a "1" if she answers "Yes," and a "0" if she answers "No" and "9" for "Don'tknow" in the box provided to the right of each vaccine dose. Do not ask the mother to supply dates from memory. Enter a date ONLY if the card is available and lists a date for the immunization dose.

**Question 1.** Record whether the child's vaccination card is present.

**Question 2.** Record date of BCG vaccination in the spaces provided, or if card is not available, record mother's response to the probing question in box at the right.

**Questions 3–5.** Record date from card in space provided, or code for mother's response to the relevant probing question in box at the right of each vaccine dose.

**Question 6.** If the BCG scar check is placed here, check the child's arm as you have been instructed, and record the result here.

### ALTERNATIVE IMMUNIZATION MODULE

**Questions 1–4:** If the child has a vaccination card, write the number of doses of each vaccine recorded on the card in the first column (headed "Card"). If no card is available, ask each question in turn, recording the number of doses the mother reports in the second column (headed "Mother's Report"). If the mother does not know whether the child received the vaccine, put a dash (—) in the space for questions 2, 3 and 4.

**Question 5:** Record result of BCG scar check here.

### ANTHROPOMETRY MODULE<sup>1</sup>

Each child is weighed, and the result recorded on his/her questionnaire. Be sure the weight for each child is recorded on the correct questionnaire, checking the child's name and line number at the top of the page.

---

<sup>1</sup>For more details on anthropometry techniques, see *Assessing the Nutritional Status of Young Children*, DP/UN/INT-88-X01/8E (New York: UN National Household Survey Capability Programme, 1990); *Demographic and Health Surveys Interviewer's Manual*, DHS-II Basic Documentation-4 (Colombia, Md.: Institute for Resource Development, 1990); and *How to Weigh and Measure Children*, DP/UN/INT-81-041-6E (New York: UN National Household Survey Capability Programme, 1986), a summary of which is included in Appendix 2 of this manual.

Follow instructions for weighing as you were trained; see Appendix 2, Anthropometric Techniques.

The mother may assist you, if necessary, when you weigh the child. Weighing (and measuring) should be done only when you have completed health questionnaires for all children in the household. That will allow you to become familiar with the household members. If there is more than one eligible child in a household, complete all weighing and measuring for that child before going on to the next child. If there is more than one mother of eligible children, weigh and measure all her children before proceeding to the next mother's children.

When you weigh and measure, you must control the child. Be firm but gentle with the child. Stay calm. The mother and the child will feel your confidence. Never leave a child alone with a piece of equipment.

Explain the weighing procedure to the mother, and to the child, to help reduce fear. Sometimes, young children are very uncooperative. You must decide if the process is causing undue stress and must be stopped. Do not weigh or measure a child if the mother refuses, or if the child is too sick or distressed.

**Question 1.** If the BCG scar record is placed here, examine the child and record here.

**Question 2.** The child should then be weighed according to the instructions given during training. Record the result to the nearest 10 grams here. Place the kilograms to the left of the decimal point and grams to the right of the decimal point.

**Question 3.** If the child is measured, record the result to the nearest tenth of a centimeter here. Also record whether the measurement was made Lying (1) or Standing (2).

**Question 4.** Record the measurer's code number.

**Question 5.** Record whether the measurement was made, or the reason it was not made.

Before you leave the household, check to ensure that the entire questionnaire is completed and that there are no blanks left. Be sure to distinguish between true zero answers and missing data or "don't know" answers.

When you have completed all the modules for each eligible child in the household, record that the questionnaire is complete for that household in the Household Form of the Household Module (page Q1).

Make sure to fill in the cluster and household number at the top of each form used for this household before leaving the dwelling. Clip the pages together as instructed and give them to your supervisor.

Thank the respondent(s) for her (their) time and cooperation.

## MORTALITY MODULE<sup>2</sup>

Use the alternative Household Module found in the Mortality Module section of the Model Questionnaire (page Q12).

First, fill out the Household Information Panel (page Q1), as before.

Second, fill out the supplementary form used to record women's birth histories (Mother Listing Form, page Q12).

List all women in the household, and their ages. Ask each woman the following questions:

1. How many children (born alive—that is, who lived and breathed or cried, even if they lived only a short time) have you given birth to who are living with you?
2. How many children (born alive—that is, who lived and breathed or cried, even if they lived only a short time) have you given birth to who are living elsewhere? (This includes grown-up children and those who have left home.)
3. How many of your children have died?
4. To check, total the responses to these three questions and ask, "Altogether, you have had \_\_\_ total children, is that right?" If the mother answers "yes," go on to the next page; if she answers "no," go back and check the three questions again.

Third, when the Mother Listing Form is complete, for each woman who has had at least one child, fill in a Pregnancy History Form (page Q13). Ask what the outcome of each pregnancy was. Record 1 if the outcome was a live birth, 2 if the birth was a stillbirth, and 3 if the pregnancy resulted in a miscarriage or abortion.

For each pregnancy that resulted in a live birth, ask each mother for the date of birth, sex, age and survival of her last, next-to-last and second-last live births, beginning the questioning with her last live birth. This procedure is used to make sure that recent-born children who died are not omitted, especially those who died shortly after birth. For children who died, record the age at death in completed months.

Finally, fill in the Child Listing Form for each mother and any other principal caretaker of children who live in this household. (The instructions are the same as for the original listing page, found on page Q1 in the Model Questionnaire.) Ensure that all children in the household whom she cares for are listed here, even if they are not her own biological children, and even if they are not at home. Stop listing when you reach a child over age 15 years.

---

<sup>2</sup>Detailed instructions for administering the mortality questionnaire are found in *Measuring Childhood Mortality: A Guide for Simple Surveys* (Amman: UNICEF Middle East and North Africa Regional Office, 1990).

## FILLING IN THE CLUSTER CONTROL SHEET<sup>3</sup>

This section is written for the supervisors and interviewers.

Select households to be included in the clusters as you have been instructed by the survey coordinator.

Instructions for filling in the cluster control sheet:

1. Record the number of eligible mothers and children in the household from the Child Listing Form in the Household Module (page Q1).
2. Record the number of completed interviews (household modules). Note on the control sheet times for re-visits to empty households, and times to interview mothers who could not be interviewed on the first visit. Keep a record of where these households are located by drawing a sketch map on the back of the page, if necessary.
3. Supervisors: Begin to check the completed questionnaires, inserting numerical codes where necessary. Look for incomplete identifications, or gaps where there should be a number. Send back questionnaires which are clearly in error.
4. When all the interviews have been completed, check that the questionnaires and the control sheet have been filled in correctly and as many mothers as possible have been interviewed. Add up the number of eligible women and children and the number of completed interviews on the control sheet. Add any comments in the "Notes" section (problems encountered, etc.).
5. Check that you have the number of questionnaires as indicated by the total completed interviews in (4) above.
6. Ensure that all the forms carry complete identification of the cluster, households, mothers and children.
7. Note any problems and discuss them with interviewers, if possible, before leaving the community.
8. Collect all completed questionnaires for one cluster together, folded in the Control Sheet.

---

<sup>3</sup>This section is adapted from "A Manual for Measurement of Childhood Mortality with Simple Surveys," developed for WHO in conjunction with UNICEF, by the Centre for Population Studies, London School of Hygiene and Tropical Medicine, June, 1990.

**CLUSTER CONTROL SHEET**

**District [name]:**

**Date: / /**

**Interviewer number: \_\_\_\_\_**

**Cluster number: \_\_\_\_\_**

**Notes:**

Household code no.	Name of head of household	Number of eligible		Interviews completed	Notes
		mothers	children		
<b>Total:</b>					

**Table A1.1 Illustrative table for converting children's dates of birth to ages****(N.B.: Must be recomputed according to actual dates of survey.)**

<b>Month and year of survey:</b>													
[7/90] [8/90] [9/90][10/90] [11/90][12/90][1/91][2/91][3/91][4/91][5/91][6/91][7/91]													
<b>Date of birth (month and year):</b>													
<b>AGE (months)</b>													
▼													
00	7/90	8/90	9/90	10/90	11/90	12/90	1/91	2/91	3/91	4/91	5/91	6/91	7/91
01	6/90	7/90	8/90	9/90	10/90	11/90	12/90	1/91	2/91	3/91	4/91	5/91	6/91
02	5/90	6/90	7/90	8/90	9/90	10/90	11/90	12/90	1/91	2/91	3/91	4/91	5/91
03	4/90	5/90	6/90	7/90	8/90	9/90	10/90	11/90	12/90	1/91	2/91	3/91	4/91
04	3/90	4/90	5/90	6/90	7/90	8/90	9/90	10/90	11/90	12/90	1/91	2/91	3/91
05	2/90	3/90	4/90	5/90	6/90	7/90	8/90	9/90	10/90	11/90	12/90	1/91	2/91
06	1/90	2/90	3/90	4/90	5/90	6/90	7/90	8/90	9/90	10/90	11/90	12/90	1/91
07	12/89	1/90	2/90	3/90	4/90	5/90	6/90	7/90	8/90	9/90	10/90	11/90	12/90
08	11/89	12/89	1/90	2/90	3/90	4/90	5/90	6/90	7/90	8/90	9/90	10/90	11/90
09	10/89	11/89	12/89	1/90	2/90	3/90	4/90	5/90	6/90	7/90	8/90	9/90	10/90
10	9/89	10/89	11/89	12/89	1/90	2/90	3/90	4/90	5/90	6/90	7/90	8/90	9/90
11	8/89	9/89	10/89	11/89	12/89	1/90	2/90	3/90	4/90	5/90	6/90	7/90	8/90
1yr	7/89	8/89	9/89	10/89	11/89	12/89	1/90	2/90	3/90	4/90	5/90	6/90	7/90
13	6/89	7/89	8/89	9/89	10/89	11/89	12/89	1/90	2/90	3/90	4/90	5/90	6/90
14	5/89	6/89	7/89	8/89	9/89	10/89	11/89	12/89	1/90	2/90	3/90	4/90	5/90
15	4/89	5/89	6/89	7/89	8/89	9/89	10/89	11/89	12/89	1/90	2/90	3/90	4/90
16	3/89	4/89	5/89	6/89	7/89	8/89	9/89	10/89	11/89	12/89	1/90	2/90	3/90
17	2/89	3/89	4/89	5/89	6/89	7/89	8/89	9/89	10/89	11/89	12/89	1/90	2/90
1yr½	1/89	2/89	3/89	4/89	5/89	6/89	7/89	8/89	9/89	10/89	11/89	12/89	1/90
19	12/88	1/89	2/89	3/89	4/89	5/89	6/89	7/89	8/89	9/89	10/89	11/89	12/89
20	11/88	12/88	1/89	2/89	3/89	4/89	5/89	6/89	7/89	8/89	9/89	10/89	11/89
21	10/88	11/88	12/88	1/89	2/89	3/89	4/89	5/89	6/89	7/89	8/89	9/89	10/89
22	9/88	10/88	11/88	12/88	1/89	2/89	3/89	4/89	5/89	6/89	7/89	8/89	9/89
23	8/88	9/88	10/88	11/88	12/88	1/89	2/89	3/89	4/89	5/89	6/89	7/89	8/89
2yr	7/88	8/88	9/88	10/88	11/88	12/88	1/89	2/89	3/89	4/89	5/89	6/89	7/89
25	6/88	7/88	8/88	9/88	10/88	11/88	12/88	1/89	2/89	3/89	4/89	5/89	6/89
26	5/88	6/88	7/88	8/88	9/88	10/88	11/88	12/88	1/89	2/89	3/89	4/89	5/89

*(Conversion Table, continued)*

27	4/88	5/88	6/88	7/88	8/88	9/88	10/88	11/88	12/88	1/89	2/89	3/89	4/89
28	3/88	4/88	5/88	6/88	7/88	8/88	9/88	10/88	11/88	12/88	1/89	2/89	3/89
29	2/88	3/88	4/88	5/88	6/88	7/88	8/88	9/88	10/88	11/88	12/88	1/89	2/89
2yr½	1/88	2/88	3/88	4/88	5/88	6/88	7/88	8/88	9/88	10/88	11/88	12/88	1/89
31	12/87	11/88	10/88	9/88	8/88	7/88	8/88	7/88	6/88	5/88	4/88	3/88	2/88
32	11/87	10/87	9/88	8/88	7/88	6/88	5/88	4/88	3/88	2/88	1/88	12/88	11/88
33	10/87	9/87	8/87	7/88	6/88	5/88	4/88	3/88	2/88	1/88	12/88	11/88	10/88
34	9/87	8/87	7/87	6/87	5/88	4/88	3/88	2/88	1/88	12/88	11/88	10/88	1/88
35	8/87	7/87	6/87	5/87	4/87	3/88	2/88	1/88	12/88	11/88	10/88	9/88	8/88
3 yr	7/87	6/87	5/87	4/87	3/87	2/87	1/88	12/88	11/88	10/88	9/88	8/88	7/88
37	6/87	5/87	4/87	3/87	2/87	1/87	12/87	11/88	10/88	9/88	8/88	7/88	6/88
38	5/87	4/87	3/87	2/87	1/87	12/87	11/87	10/87	9/88	8/88	7/88	6/88	5/88
39	4/87	3/87	2/87	1/87	12/87	11/87	10/87	9/87	8/87	7/88	6/88	5/88	4/88
40	3/87	2/87	1/87	12/87	11/87	10/87	9/87	8/87	7/87	6/87	5/88	4/88	3/88
41	2/87	1/87	12/87	11/87	10/87	9/87	8/87	7/87	6/87	5/87	4/87	3/88	2/88
3yr½	1/87	12/87	11/87	10/87	9/87	8/87	7/87	6/87	5/87	4/87	3/87	2/87	1/88
43	12/86	11/87	10/87	9/87	8/87	7/87	6/87	5/87	4/87	3/87	2/87	1/87	12/87
44	11/86	10/86	9/87	8/87	7/87	6/87	5/87	4/87	3/87	2/87	1/87	12/87	11/87
45	10/86	9/86	8/86	7/87	6/87	5/87	4/87	3/87	2/87	1/87	12/87	11/87	10/87
46	9/86	8/86	7/86	6/86	5/87	4/87	3/87	2/87	1/87	12/87	11/87	10/87	9/87
47	8/86	7/86	6/86	5/86	4/86	3/87	2/87	1/87	12/87	11/87	10/87	9/87	8/87
4 yr	7/86	6/86	5/86	4/86	3/86	2/86	1/87	12/87	11/87	10/87	9/87	8/87	7/87
49	6/86	5/86	4/86	3/86	2/86	1/86	12/86	11/87	10/87	9/87	8/87	7/87	6/87
50	5/86	4/86	3/86	2/86	1/86	12/86	11/86	10/86	9/87	8/87	7/87	6/87	5/87
51	4/86	3/86	2/86	1/86	12/86	11/86	10/86	9/86	8/86	7/87	6/87	5/87	4/87
52	3/86	2/86	1/86	12/86	11/86	10/86	9/86	8/86	7/86	6/86	5/87	4/87	3/87
53	2/86	1/86	12/86	11/86	10/86	9/86	8/86	7/86	6/86	5/86	4/86	3/87	2/87
4yr½	1/86	12/86	11/86	10/86	9/86	8/86	7/86	6/86	5/86	4/86	3/86	2/86	1/87
55	12/85	11/86	10/86	9/86	8/86	7/86	6/86	5/86	4/86	3/86	2/86	1/86	12/86
56	11/85	10/85	9/86	8/86	7/86	6/86	5/86	4/86	3/86	2/86	1/86	12/86	11/86
57	10/85	9/85	8/85	7/86	6/86	5/86	4/86	3/86	2/86	1/86	12/86	11/86	10/86
58	9/85	8/85	7/85	6/85	5/86	4/86	3/86	2/86	1/86	12/86	11/86	10/86	9/86
59*	8/85	7/85	6/85	5/85	4/85	3/86	2/86	1/86	12/86	11/86	10/86	9/86	8/86

# APPENDIX 3





## Organization and Administration of Data Processing

### CONTENTS

Managing the Data Base .....	A3.4
Record-Keeping .....	A3.4
How to Install and Use the Data Entry and Analysis Programs .....	A3.6
Epi Info Package for Mid-Decade Goal Surveys .....	A3.6
Installing the Programs .....	A3.6
Running the Program .....	A3.6
Data Entry .....	A3.7
Notes on the specific modules .....	A3.8
Consistency Checks .....	A3.8
Variable Frequencies .....	A3.8
Data Analysis .....	A3.8
Data Back-Up .....	A3.9
Enter EPI INFO .....	A3.9
Quit to DOS Sub-directory .....	A3.9
Running CSAMPLE to Obtain Standard Errors for Cluster Surveys .....	A3.9
Notes to the Programmer .....	A3.11
Removing Modules .....	A3.11
Adding Modules .....	A3.11
Changing the Wording of Questions or Response Categories .....	A3.11
Changing the Analysis/Consistency Checks .....	A3.12
Merging Data Sets from Several Computers .....	A3.12

## MANAGING THE DATA BASE

We suggest that you carry out data entry and data editing in small batches, usually completing data entry for an entire cluster at once. Data should be stored on the hard disk. After a batch has been entered, the data are passed through a consistency checking program and the results are printed out.

Any errors which are due to incorrect keying should immediately be corrected. The list of errors should be given to the supervisor to decide what action should be taken (for example, changing the data, consulting with the interviewer or leaving the error unresolved). Once corrections for errors have been made to the data, the consistency program should be rerun to verify the corrections. When there are no more errors that can be corrected in the office, the data should be backed up to diskette.

☞ Label your diskettes with the name of the files, the region and the date of the latest changes.

When the complete data base is ready for analysis, the checking program should be used again as a final check before frequencies are produced for each question.

## RECORD-KEEPING

Data processing operations frequently become messy because of the volume of data and the multiple procedures which have to be carried out. Good record-keeping can alleviate these problems.

Figure A3.1 shows a sample record-keeping form. This form has a line for each batch or cluster of data.

1. Record the numbers of questionnaires that return to the office from the field for each cluster. Check these numbers against the field work totals.
2. When a batch of questionnaires is submitted for data entry, record the date.
3. When a batch has been entered, record the date, the number of questionnaires and the number with errors.
4. Note the date that corrections are made.
5. Each time error reports from a batch are produced, note the number of documents with errors and the date.
6. When the main database is updated with the batch, note the date and number of records added.

Data from the questionnaire are entered for each household in the order they are collected: the household data from the listing page, the data on all children on the child roster, the water and sanitation and salt iodization modules, education data for each child over school-entry age, tetanus toxoid module for each mother, care of acute respiratory illness (CARI) module for each mother and, last, child health modules for each child. Once all the data for one household are entered, the clerk moves on to the next household.



The data entry can be done by a number of data entry operators. Error reports are produced at the end of each cluster's data entry with a description of any inconsistent items.

Corrections, written on the error sheets, are used to correct the appropriate records on diskette. When no more errors that can be corrected are encountered, the data should be backed up to diskette.

## **HOW TO INSTALL AND USE THE DATA ENTRY AND ANALYSIS PROGRAMS**

### **EPI INFO PACKAGE FOR MID-DECADE GOAL SURVEYS:**

#### **INSTALLING THE PROGRAMS**

All the files for data entry and analysis should be placed in a separate directory on the hard drive. To install the software, first create a new directory on the hard drive by typing at a DOS prompt:

```
MD C:\MIDDECAD
```

Then copy all of the files on the diskette into this directory by typing:

```
COPY A:*. * C:\MIDDECAD
```

It is assumed that EPI INFO v.6 is already installed on the hard drive. If not, follow the procedures described in the EPI INFO manual for installation. The \EPI6 directory should be included in the computer's PATH. You can verify that it is by typing PATH at a DOS prompt. If \EPI6 is not included in the PATH, modify your AUTOEXEC.BAT file to include it.

#### **RUNNING THE PROGRAM**

The software should always be run from the \MIDDECAD directory. First switch to this directory by typing:

```
C:  
CD \MIDDECAD
```

Then to run the programs, type:

```
MENU
```

This will bring up a menu of choices including:

1. Data Entry
2. Consistency Check
3. Variable Frequencies
4. Data Analysis
- B. Data Backup
- Q. Quit to DOS Sub-directory

## Data Entry

Choose option 1 from the menu to either enter new data or to edit already existing data. Choosing this option will bring up the first data-entry screen for entering data from the household information sheet. A new questionnaire may be immediately input. To edit an existing questionnaire, press <CTRL F> and enter the CLUSTER and HOUSEHOLD numbers of the desired questionnaire. Then press <F3> and <ENTER> to bring up this questionnaire for editing. Pressing <CTRL N> will again allow new questionnaires to be input. For further details on moving about the data-entry screens within EPI INFO, refer to the EPI INFO manual on data entry.

Data entry is organized into several modules, including:

1. Household information (single occurrence)
2. Mother and child roster (multiple occurrences)
3. Water and sanitation (single occurrence)
4. Salt iodization (single occurrence)
5. Education (multiple occurrences)
6. Tetanus toxoid (multiple occurrences)
7. Care of acute respiratory infection (multiple occurrences)
8. Child health (multiple occurrences)

This organization allows the data-entry staff to complete all data entry for a page of the questionnaire before moving on to the next page, and thus to eliminate flipping back through the questionnaire.

Entry into each module is controlled at the bottom of the household screen. To begin entering data from the mother and child roster, simply press <ENTER> next to GoToChild. This will bring up a screen to enter the first line in the roster. When the data are entered and written to disk, the same screen is repeated, since multiple lines are allowed. After entering all lines from the roster, press <F10> (RETURN) and the household screen will reappear. Press <ENTER> next to GoToWater to go to the Water and Sanitation Module. Since only a single occurrence of the Water and Sanitation Module is allowed for each household, the household screen is immediately reentered upon saving the water record. The same procedures are applied for the remaining modules. If there are no data for a module, the module can be skipped by moving the cursor beyond the module before pressing <ENTER> or by simply pressing <F10> immediately upon entering the module. When data have been entered for all modules, pressing <ENTER> next to Finish will end the data entry.

The cluster and household numbers are concatenated to form a HOUSEID which serves as a key to link all of the files together. The HOUSEID must uniquely identify the household and cannot be repeated. Additionally, the mother's or child's line number serves as a key to link the individual's information to the appropriate line in the mother and child roster. It is important that this line number be entered correctly as many of the indicators rely on linking data from the roster and the specific modules.

**Notes on the specific modules:**

*Household information panel:* The date of interview must be entered because all ages and dates of birth must be verified against the date of interview.

*Mother and child listing form:* Line numbers for mothers or caretakers must end in a 0 (i.e., 1-0, 2-0, etc.). Date of birth is not entered for them. Line numbers for children must end in a number from 1 to 9 (i.e., 1-1, 1-2, etc.)

*Tetanus toxoid:* The mother's line number (ending in 0) must be entered here to be used to link the mother to her youngest child. (The key for linking children is based on the line number plus 1, since dates of immunization are to be compared against the youngest child's date of birth.) Data for question 6 need not be entered, since the answer to question 6 is simply a total of questions 2 and 4.

*CARI:* The mother's line number (ending in 0) must be entered, although the key for linking is based on the line number plus 1.

*Children under age five years of age:* Because all of the modules for children under five years of age (diarrhoea, vitamin A, breastfeeding, immunization, and anthropometry) are to be filled out for one child before going on to the next child, they are treated as a single module for data entry. Data on the date of interview and the child's date of birth must be reentered in this module because EPI INFO will calculate anthropometric indices at the time of data entry.

**Consistency Checks**

Data for each cluster should be checked for internal consistency as soon as all the data have been entered for the cluster. Choose option 2 to perform consistency checking. You will be prompted as to whether you want to send the results directly to the printer or to a file for future examination. You will also be prompted as to what range of clusters are to be checked. If you are only interested in cluster 101, type 101 as the starting and as the ending cluster numbers. If you would like to run consistency checks for the entire file, type 000 as the starting cluster and 999 as the ending cluster.

Consistency errors will be written to a file called CHECK.TXT or to the printer directly. Each error statement shows: (1) the household identification number, (2) the questionnaire module that contained the error, (3) a description of the error and (4) the individual number, if applicable.

**Variable Frequencies**

When all the data have been entered, it will be useful to examine a frequency listing of all the variables in the data file. Choosing option 3 will produce a frequency listing of all the variables with the exception of HOUSEID and CHILDDID.

**Data Analysis**

Choose option 4 to analyse the data. This option runs the set of commands contained in INDICATR.PGM, which produces the results for the indicators, according to the specifications laid out in Appendix 4. You can obtain a full listing of the .PGM files by simply printing the file, either to the screen or to your printer. The

output is limited to looking at the specific indicators for the entire sample. The program can be modified to look only at specific subgroups of the population, and then rerun for that subgroup. Instructions on modifying the program are given in the section "Notes to the Programmer" below.

The analysis program does not supply standard errors for the indicators. These may be calculated using the EPI INFO v.6 CSAMPLE option. See below for instructions on how to use CSAMPLE.

### **Data Back-Up**

Data should be backed up to diskette on a daily basis, using different diskettes on different days of the week. Choosing option B will copy the appropriate files to the diskette. This procedure assumes that the size of the data files will not exceed the capacity of a single diskette. In the event that the data files become too large for a diskette, the data manager will need to find an alternative system for backing up and modify the MENU.BAT file accordingly. Systems which might be considered include PKZIP, BACKUP (DOS), or PCTOOLS.

### **Enter EPI INFO**

Should there be a need to work directly with the EPI INFO package, choosing Option E will take you to the EPI INFO v.6 startup screen. All the EPI INFO programs and facilities can be run from this screen. When exiting EPI INFO, you will return to the menu. (Option E does not actually appear on the menu list, since data-entry staff will not regularly need to enter EPI INFO directly. The option is functional, however.)

### **Quit to DOS Sub-directory**

Choose option Q to exit the menu system.

### **RUNNING CSAMPLE TO OBTAIN STANDARD ERRORS FOR CLUSTER SURVEYS**

To estimate correct standard errors for the indicators in EPI INFO, a data set must be on disk that has the created variables for the indicator of interest. For example, suppose we are interested in obtaining standard errors for Indicator 11.1, the per cent of under-fives with low weight-for-age. To create such a data set, add two lines to the INDICATR.PGM file after the line "report waz.rpt":

```
ROUTE TEMP.REC  
WRITE "RECFILE" CLUSTER LOW2WAZ
```

and run the analysis. This will create a temporary file called TEMP.REC, which contains only the cluster variable and the variable on whether or not the child has low weight-for-age.

Then choose the CSAMPLE option from the PROGRAMS menu option in EPI INFO v.6. Type TEMP as the file name. A screen of options will appear. Under Main, type LOW2WAZ, the variable for which you need standard errors. Specify the strata variable only if the sample design was stratified. Specify the PSU (in chapter 4, the term "small area" refers to these Primary Sampling Units) variable, usually given the name CLUSTER. Weight is not needed unless the sample design requires weights. Results may be viewed immediately. Specify a Crosstab variable only if you need to examine standard errors within population



subgroups. ly on the screen, printed directly to a printer or stored in a file for future use. If the data are not already sorted by CLUSTER, then choose SORT. Choose TABLES, and the analysis will be performed.

The output gives the standard error for the indicator and its 95 per cent upper and lower confidence intervals. The program also computes the design effect obtained in the survey for the variable of interest. For instance, in analysing weight-for-age, the output might look like the example shown in Figure A3.2, in which the percent (rounded to the nearest tenth) with low weight-for-age is 21.0 with a confidence interval of (19.3, 22.7).

**Figure A3.2 Example of CSAMPLE output when analysing weight-for-age**

**CTABLES COMPLEX SAMPLE DESIGN ANALYSIS**

Analysis of LOW2WAZ  
LOW2WAZ

	Total
0	
Obs	2168
Percent V	79.009
SE%	0.887
LCL%	77.271
UCL%	80.746
1	
Obs	576
Percent V	20.991
SE%	0.887
LCL%	19.254
UCL%	22.729
Total Obs	2744
Design eff.	1.300

LCL = lower confidence limit  
UCL = upper confidence limit

Sample Design Included:

-----  
Sampling Weights--None  
Primary Sampling Units from CLUSTER  
Stratification--None

0 records with missing values

## NOTES TO THE PROGRAMMER

Some modifications to the programs will be required in every country where the package is used. For example, the definition of "safe and convenient water" is country-specific, and therefore, the analysis program INDICATR.PGM must be edited to reflect the definition used in a specific country. The definition of school-entry age varies from country to country. Also, alternative modules exist for vitamin A as well as for immunization coverage. A programmer who is familiar with EPI INFO must make the appropriate country-specific modifications to the program. It is best that the programmer fully understand the structure of the data base prior to making any modifications. This is especially true if there are to be many changes to the questionnaire.

Any changes to a .QES or .CHK file require that the corresponding .REC file be recreated. This can be accomplished by choosing ENTER in EPI INFO v.6 and choosing option 2, "Create new data file from .QES file." Since the data are contained in the .REC file, it is important that all changes to .QES and .CHK files be made before data entry begins. If it is necessary to make some modification to these files after data have been entered, consult the EPI INFO manual on "Revising the Structure of a Data File (Menu Choice 3)," in chapter 8, "Entering Data."

### Removing Modules

Any modules in the sample questionnaire that a country chooses not to implement can easily be skipped. If the module to be deleted is one of those listed above on page A3.6, a simple deletion in the HOUSEHLD.CHK file will prevent ENTERING the module. The variable in the HOUSEHLD.QES for that module should also be deleted.

If only a portion of a data-entry module is to be deleted (for example, the breastfeeding questions within the child health module), then the lines referring to those questions should be deleted from the .QES file and the .CHK file. The alternative vitamin A and immunization modules which are not implemented should be removed in this fashion. The appropriate number of blank lines should be added or deleted to the .QES file to ensure that each page breaks on the screen at a logical place.

### Adding Modules

If additional modules are to be added to the questionnaire, it will be necessary to develop new .QES and .CHK files. The existing files can serve as templates of how to do this. In most cases, additional questions could be added to the existing modules, in which case, modification of the .QES and .CHK files should suffice.

### Changing the Wording of Questions or Response Categories

Changes to the wording of questions (including translation into languages other than English) simply requires that the screens defined in the .QES files be modified to show the appropriate wording. The original variable names which are used in the .QES, .CHK and .PGM files need not be changed. However, if the variable names are changed, the changes must be made in all of these files.

If the categories of acceptable responses to questions are changed, appropriate changes must be made to the .CHK file (to allow all valid responses but not allow invalid responses) and to the .PGM file (to ensure that the indicators are calculated using the correct categories). The package does not explicitly assign labels to the responses, so that there is generally no need to type in labels to the responses.

### Changing the Analysis/Consistency Checks

Both consistency checking and calculation of the indicators are accomplished through .PGM files. These files can be edited to remove references to any modules not implemented and to specify the definition of certain indicators (e.g., safe and convenient water, school-entry age and frequency of vitamin A capsule use). The indicators are usually desired for other specific subgroups, such as region-specific indicators. The indicators are usually desired for other specific subgroups, such as region-specific, or sex-specific, indicators.

- ☞ Obtaining estimates for sub-groups: To calculate each indicator for specific subsets of the data, the easiest modification would be to select only those observations in the subset, and rerun the analysis program. For example, to calculate the indicators for males only, select the male observations:

```
SELECT SEX = 1 .
```

Since the select statement must be executed every time a new data set is read, the INDICTR.PGM includes a subroutine at the end called ":SELECTSEX" to perform the analysis for whatever subsets of the data are required. Edit this subroutine and then run the revised program to calculate the indicators for boys only. For girls only, revise the subroutine with the line:

```
SELECT SEX = 2
```

and rerun the INDICATR.PGM file. The select statement can also be revised to consider only observations in a specific REGION, AGE group or URBAN/RURAL category. Of course, the programmer should make sure that the variable of interest is always available by issuing an appropriate RELATE command before the selection.

- ☞ Additional .PGM files can be created for other purposes—for example, to compute the mean duration of breastfeeding or to examine the age distribution of children. Examination of the existing .PGM files can be helpful in getting started on new analysis.
- ☞ All the EPI INFO analysis programs are written assuming that the sample in the country is designed to be self-weighting (i.e., the weights are all 1.0). However, a subroutine is included to make weighted analysis as straightforward as possible. The programmer will simply need to edit the subroutine ":WEIGHT" at the end of the INDICATR.PGM file, to include assignment of weights as described in chapter 7.

**Merging Data Sets from Several Computers**

In some cases, it will be necessary to enter questionnaire data on several computers and then merge the data together for analysis. This can be accomplished through the MERGE program in EPI INFO. The .REC files from one computer should be copied onto the other, but placed in a separate subdirectory. The MERGE program, using the CONCATENATE option, will create a new data file with data from both files. The procedure can be repeated if more than two computers are used. The MERGE must be completed separately for each module.

# APPENDIX 4



## DEFINITIONS OF THE INDICATORS

Indicator	Numerator	Denominator
1.1 BCG coverage	Number 12–23-month-olds receiving BCG vaccine before 1st birthday <sup>1</sup>	Total number of 12–23-month-olds surveyed
1.2 DPT coverage	Number 12–23-month-olds receiving DPT3 vaccine before 1st birthday <sup>2</sup>	Total number of 12–23-month-olds surveyed
1.3 OPV coverage	Number 12–23-month-olds receiving OPV3 vaccine before 1st birthday	Total number of 12–23-month-olds surveyed
1.4 Measles coverage	Number 12–23-month-olds receiving measles vaccine before 1st birthday	Total number of 12–23-month-olds surveyed
1.6 Tetanus toxoid (TT) vaccine coverage	Number of mothers of 0–11-month-olds with at least two doses of TT within 3 years of child's birth <sup>3</sup>	Total number of mothers of under-one-year-olds surveyed
5.1 Vitamin A Supplementation Programme	Number of under-twos receiving supplement with correct timing of last dose	Total number of under-twos surveyed
5.1 Vitamin A Fortification Programme	Number of children under 2 with fortified food product in home AND who received the food in last week	Total number of under-twos surveyed

<sup>1</sup>For “valid” vaccine coverage, only children with immunization dates for BCG, DPT3, OPV3 or Measles occurring before the first birthday and verified by card are included in the numerator. “Valid” coverage cannot be obtained if the Alternative Immunization Module is used. For “card only” coverage, the number of 12–23-month-olds receiving BCG, DPT3, OPV3 or Measles vaccine by the time of the survey (verified by a card entry) are included in the numerator. For “card + history” coverage, number of 12–23-month-olds vaccinated with BCG, DPT3, OPV3 or measles vaccine, either entered on the card or reported by the mother, are included in the numerator; the denominator for “card + history” coverage excludes those children whose mothers answered “don't know” to the questions.

<sup>2</sup>If the Alternative Immunization Module is used, this is the number of children receiving at least three DPT (or three OPV) immunizations reported by the mother. All 12–23-month-old children form the denominators, but exclude children whose mothers reply “don't know” to the questions.

<sup>3</sup> Numerator is all mothers of under-ones with at least two TT doses during last pregnancy OR with last dose within appropriate interval for number of reported doses in lifetime.

Indicator	Numerator	Denominator
5.1 Vitamin A Dietary Education Programme	Number of under-twos whose mothers have heard programme message AND give "correct" response to foods that message promotes (knowledge); and Number of under-twos whose mothers have heard programme message AND who received the food in the last week (practice)	Total number of under-twos surveyed
6.1 Salt iodization	Number of households with salt testing positive for iodine/iodate	Total number of households surveyed <sup>4</sup>
7.1 ORT use (pre-1993 definition)	Number of diarrhoea cases among under-fives in 2 weeks before survey who received ORT and/or recommended home fluids <sup>5</sup>	Total number of diarrhoea cases <sup>6</sup> among under-fives in two weeks preceding survey
7.2 ORT use (increased fluids and continued feeding)	Number of diarrhoea cases among under-fives taking "more" fluids <sup>7</sup> AND continued eating somewhat less, the same or more food	Total number of diarrhoea cases <sup>8</sup> among under-fives in two weeks preceding survey
11.1, 11.2 Nutritional status: weight-for-age	Number of under-fives who fall below -2 standard deviations from the median weight-for-age of the NCHS/WHO standard; number who fall below -3 SDs	Total number of under-fives <sup>9</sup> weighed
11.3, 11.4 (OPTIONAL) Nutritional status: height-for-age	Number of under-fives who fall below -2 standard deviations from the median height-for-age of the NCHS/WHO standard; number who fall below -3 SDs	Total number of under-fives <sup>10</sup> measured

<sup>4</sup>If a household has salt, but it is not tested (Q.1 = 9), these households are omitted from the denominator.

<sup>5</sup>Q.2A, B, C, D, E, F = 1. That is, Q. 2G, 2H, 2I NOT equal to 1 or 9.

<sup>6</sup>If Q. 1 = 9 (don't know if child has had diarrhoea in past two weeks), the child is omitted from the calculation.

<sup>7</sup>"More" fluids here refers to ANY fluid noted in Q.2, whether or not it is classified as "recommended fluid."

<sup>8</sup>See footnote 6.

<sup>9</sup>Children with out-of-range weights-for-age are omitted from calculations.

<sup>10</sup>Children with out-of-range heights-for-age are omitted from calculations.



Indicator	Numerator	Denominator
12.1 Education: retention rates to grade 5	Proportion of children entering first grade of primary school who eventually reach grade 5 <sup>11</sup>	All children still enrolled in primary school who attended school during previous year <sup>12</sup>
12.2 Education: net enrollment rate	Number of children currently enrolled in primary school of primary-school age <sup>13</sup>	Total number of children of primary-school age surveyed
12.3 Education: primary-school entry rate	Number of children of school-entry age who are currently attending grade 1 <sup>14</sup>	Total number of children of school-entry age surveyed
12.4 Education: gross enrollment rate	Number of under-15s currently enrolled in primary school <sup>15</sup>	Total number of children of primary-school age surveyed
13.1 Water supply	Number of household residents <sup>16</sup> in defined “safe and convenient” categories	Total number of household residents <sup>17</sup> surveyed
13.2 Sanitation	Number of household residents in defined “safe and convenient” categories <sup>18</sup>	Total number of household residents <sup>19</sup> surveyed
OPTIONAL Indicator: Exclusive Breastfeeding	Number of infants less than 4 months of age exclusively breastfed <sup>20</sup>	Total number of infants 0–3 months of age surveyed

<sup>11</sup>The EPI INFO program uses Q.4 and Q.5 to calculate transition probabilities for the cohort of children in the sample to estimate this indicator. See programs accompanying this handbook.

<sup>12</sup>Please see footnote 11.

<sup>13</sup>Select children in primary school age range; Education Module Q.2= 1.

<sup>14</sup>Select children of primary school entry age; Education Module Q.2=1 and Q.3 Grade=1, Level=1.

<sup>15</sup>Number of under-15-year-olds with Q.2 =1 and Q.3 Level =1.

<sup>16</sup> This indicator is obtained by weighting number of households in this category by the number of residents in the household.

<sup>17</sup>To express this indicator in absolute numbers, the proportion of the sample population with “safe and convenient” water supply must be multiplied by the total population size. Households with missing data for this indicator are omitted from calculation.

<sup>18</sup>See footnote 16.

<sup>19</sup>See footnote 17.

<sup>20</sup>Numbers still breastfed (Q.2 = 1) and no other food given (Q.3H = 1).

<b>Indicator</b>	<b>Numerator</b>	<b>Denominator</b>
OPTIONAL Indicator: Timely complementary feeding	Number of children receiving breast milk and complementary foods <sup>21</sup>	Total number of infants 6–9 months old surveyed
OPTIONAL Indicator: Continued breastfeeding at 1 year, 2 years	Number infants 12–15 months, 20–23 months still breastfeeding <sup>22</sup>	Total number of 12–15-month-olds, 20–23-month-olds surveyed
OPTIONAL Indicator: Bottle-feeding rate	Number of infants less than 12 months old receiving any food or drink from a bottle <sup>23</sup>	Total number of under-ones surveyed
OPTIONAL Indicator: Ever breastfed rate	Number of infants less than 12 months old ever breastfed <sup>24</sup>	Total number of under-ones surveyed
OPTIONAL Indicator: Control of Acute Respiratory Infections	Number of mothers of under-fives who know signs of ARI <sup>25</sup>	Total number of mothers of under-fives surveyed

<sup>21</sup>Numbers still breastfed (Q.2 = 1) and complementary foods given last 24 hours (Q.3G =1), even if also given other breast milk substitutes.

<sup>22</sup>Numbers still breastfeeding (Q.2 = 1).

<sup>23</sup>Q.4 = 1.

<sup>24</sup>Q.1 = 1.

<sup>25</sup>Mothers answering CARI Module, Q.1 = 4 or 5, irrespective of other answers.