WHO/IVB/04.23
ORIGINAL: ENGLISH

Immunization coverage cluster survey – Reference manual

Immunization, Vaccines and Biologicals

World Health Organization

The World Health Organization has managed cooperation with its Member States and provided technical support in the field of vaccine-preventable diseases since 1975. In 2003, the office carrying out this function was renamed the WHO Department of Immunization, Vaccines and Biologicals.

The Department's goal is the achievement of a world in which all people at risk are protected against vaccine-preventable diseases. Work towards this goal can be visualized as occurring along a continuum. The range of activities spans from research, development and evaluation of vaccines to implementation and evaluation of immunization programmes in countries.

WHO facilitates and coordinates research and development on new vaccines and immunization-related technologies for viral, bacterial and parasitic diseases. Existing life-saving vaccines are further improved and new vaccines targeted at public health crises, such as HIV/AIDS and SARS, are discovered and tested (Initiative for Vaccine Research).

The quality and safety of vaccines and other biological medicines is ensured through the development and establishment of global norms and standards (Quality Assurance and Safety of Biologicals).

The evaluation of the impact of vaccine-preventable diseases informs decisions to introduce new vaccines. Optimal strategies and activities for reducing morbidity and mortality through the use of vaccines are implemented (Vaccine Assessment and Monitoring).

Efforts are directed towards reducing financial and technical barriers to the introduction of new and established vaccines and immunization-related technologies (Access to Technologies).

Under the guidance of its Member States, WHO, in conjunction with outside world experts, develops and promotes policies and strategies to maximize the use and delivery of vaccines of public health importance. Countries are supported so that they acquire the technical and managerial skills, competence and infrastructure needed to achieve disease control and/or elimination and eradication objectives (Expanded Programme on Immunization).

World Health Organization

Department of Immunization, Vaccines and Biologicals
Family and Community Health

World Health Organization
CH-1211 Geneva 27
Switzerland
Fax: +41 22 791 4227
Email: vaccines@who.int
or visit our web site at: http://www.who.int/vaccines-documents

WHO/IVB/04.23
ORIGINAL: ENGLISH

Immunization coverage cluster survey – Reference manual

Immunization, Vaccines and Biologicals

World Health Organization

The Department of Immunization, Vaccines and Biologicals thanks the donors whose unspecified financial support has made the production of this publication possible and also gratefully acknowledges the assistance given to the authors by:

Steve LWANGA, Steve SAPIRIE, Robert STEINGLASS, George STROH, Alisdair WYLIE

This publication was produced by the
Vaccine Assessment and Monitoring Team
of the Department of Immunization, Vaccines and Biologicals

Ordering code: WHO/IVB/04.23
Printed: June 2005

This publication is available on the Internet at:
www.who.int/vaccines-documents/

Copies may be requested from:
World Health Organization
Department of Immunization, Vaccines and Biologicals
CH-1211 Geneva 27, Switzerland
• *Fax:* + 41 22 791 4227 • *Email:* vaccines@who.int •

Printed by the WHO Document Production Services, Geneva, Switzerland

Contents

List of boxes figures, forms and tables

Boxes

Figures

Forms

Tables

Abbreviations

ANC	antenatal care
DEFF	design effect
DTP	diphtheria–tetanus–pertussis (vaccine)
HC	health centre
HepB	hepatitis B (vaccine)
Hib	*Haemophilus influenzae* type b (vaccine)
HOS	hospital (dose received in a hospital)
MMR	measles-mumps-rubella vaccine*
MR	measles-rubella vaccine*
NGO	non-governmental organization
NIDs	national immunization days
OPV	oral polio vaccine
OTH	other (dose received on any occasion other than those listed)
OUT	outreach
PPS	probability proportional to size (sampling)
PRIV	private (dose received from a private practitioner)
PSUs	primary sampling units
SCH	school (dose received during a school immunization campaign)
SIA	supplementary immunization activity/activities

* Measles vaccination is given either as a single antigen (measles only), or in combination with rubella (MR) or with rubella and mumps (MMR). This survey manual uses "measles vaccine" to indicate measles-containing vaccine. In countries where MMR or MR is used, the manual should be adapted accordingly.

TBA	traditional birth assistant
Td	tetanus and diphtheria toxoid – adult dose (vaccine)
TT	tetanus toxoid (vaccine)
WCV	well-child visit
YF	yellow fever (vaccine)

Glossary

1-sided test	A statistical test when the difference tested is directionally specified beforehand. For example, testing whether immunization coverage is higher in one area than in another.
2-sided test	A statistical test when the difference tested is not directionally specified beforehand. For example, testing whether immunization coverage rates in two areas are different.
Cluster	A collection of elements (for example, households, communities, villages, etc.) grouped within defined geographical or administrative boundaries.
Cluster survey	A survey in which, after the population under study has been subdivided into clusters, only (some) subjects from selected clusters are observed.
Confidence level	A level of confidence set in computing confidence limits. A level of 95% (or 0.95) is conventionally used but can be set higher or lower. A level of confidence of 95% implies that 19 out of 20 times the results from a survey using these methods will capture the true population value.
Confidence limits	The upper and lower limits of the confidence interval in interval estimation. The interval itself is called the confidence interval or confidence range. Confidence limits are so called because they are determined in accordance with a specified or conventional level of confidence or probability that these limits will in fact include the population parameter being estimated. Thus, 95% confidence limits are values between which we are 95% confident that the population parameter being estimated will lie. Confidence limits are often derived from the standard error (SE).
Design effect	A measure of variability due to selection of survey subjects by any method other than simple random sampling. It is defined as the ratio of the variance with other types of sampling to the variance with simple random sampling. Usually, cluster surveys have a design effect greater than one (the variability is higher than for simple random sampling).

Immunization cluster sampling technique	A household-based survey done in a number of clusters of a predetermined number of children to assess immunization services.
Fully immunized child (FIC)	Usually, this is a child who has received doses of the "standard six" antigens – BCG, diphtheria–tetanus–pertussis (DTP) (3 doses), polio (3 doses), and measles vaccines. In countries at risk for yellow fever, this should be included. New vaccines (hepatitis B, and *Haemophilus influenzae* type b [Hib]) are not usually included in this definition; in some countries, BCG is excluded from this definition. The definition of FIC used in a survey should be specified.
Household	A group of persons who live and eat together. This definition groups people most likely to be similarly influenced in the way they take care of their children.
Immunization coverage	Proportion of individuals in the target population who are vaccinated.
Immunization coverage target	A goal that is prepared for a health facility that states what proportion of individuals in the target population will be vaccinated with specific vaccines in a given time period.
Infant immunization coverage	The percentage of children under one year of age who have been vaccinated.
Morbidity	Illness
Mortality	Death
National immunization days (NIDs)	Supplementary mass immunization activities that complement routine immunization, and are usually carried out within a short time period. The aim of mass campaigns is to target **every** person in the defined target group, regardless of prior immunization status.
Protection at birth (PAB)	The proportion of children who are protected at birth against tetanus thanks to an up-to-date Td (or TT) immunization status of their mother.
Power (of a statistical test)	Ability to reject the test hypothesis when it is false.
Random number	A number selected by chance.
Reasons for failure to immunize	Reasons why children and women do not come or do not return for immunization.

Sampling frame	The set of sampling units from which a sample is to be selected; for example, a list of names, or places, or other items to be used as a sampling unit.
Sampling unit	The unit of selection in the sampling process; for example, a child in a household, a household in a village or a district in a country. It is not necessarily the unit of observation or study.
Simple random sampling	Selection of study subjects with equal probability of selection for each subject.
Statistical significance	The concept by which results are judged as being due to chance or not.
Supplementary immunization activity/ activities (SIA)	Any immunization activity conducted in addition to the provision of routine immunization services.
Target population	Group of individuals who are included in the immunization services based on their age, sex, and the area in which they live.
Tetanus toxoid immunization of pregnant women*	Whether mother of infants has been immunized with sufficient recent doses of tetanus toxoid (Td [or TT]), indicating that her children are protected against neonatal tetanus at birth.
TT2+ coverage**	The proportion (%) of women who have received their second or higher dose of tetanus toxoid-containing vaccine (Td [or TT]) during the most recent pregnancy.
Valid doses	Doses that were administered when the child had reached the minimum age for the vaccine, and were administered with the proper spacing between doses according the national schedule.

* WHO recommends the use of Td instead of TT. This survey manual uses "Td (or TT)" to indicate tetanus toxoid-containing vaccine. Countries should use the terminolgy (TT or Td) that is appropriate in their situation.

** To indicate coverage levels with tetanus toxoid-containing vaccine the terms "TT1, TT2, TT3, TT4, TT5, TT2+" are maintained throughout this manual.

1. Introduction

1.1 Background

Immunization staff have used the *Expanded Programme on Immunization Coverage Survey Manual* (WHO/EPI/MLM/91.10) extensively over the past 10 years as a "blueprint" for conducting immunization coverage surveys. The manual provides a prescriptive approach to the coverage survey by specifying a sample of seven children from each of 30 clusters. It provides guidance for identifying a starting household and subsequent households using what were considered simple methods that could be easily followed. The original Expanded Programme on Immunization (EPI) coverage survey was designed based on the assumption that immunization coverage was 50% to allow for maximum sample size with a precision of ±10%, in line with the low coverage levels at the time.

During the 1980s and 1990s many countries were interested in estimating their levels of immunization coverage either nationally or subnationally to establish baseline information or to satisfy demands of the partner agencies. The methods and the manual satisfied the survey objectives in the early years of immunization programmes. Many countries did not have information on the level of immunization; hence the assumption of 50% was acceptable. A precision level of ±10% was also tolerable since immunization coverage was expected to be fairly low. Instructions to the field teams, such as spinning a bottle on even ground in a "central location in the village or town, such as a market, a mosque or church", were recommended on the assumption that sampling frames could not be established.

Immunization activities in all countries have progressed over the years. In many countries, high levels of immunization coverage are not being achieved and immunization staff often wish to assess *changes over time* or *differences between geographical or administrative areas*. These needs could not be satisfied by the prescriptions of the *EPI Coverage Survey Manual*. The current revised version is less prescriptive and provides additional information on alternative methods that could be used for immunization coverage surveys. This revised manual provides guidance on conducting high quality cluster surveys for measuring levels and validity of immunization coverage and reasons for non-immunization.

1.2 The changes and structure

The user of the manual will note the following major changes from the 1991 version:

- methods for identifying the starting household without "spinning bottles in the market place" in Section 3.5.3;
- methods for identifying the "next household" in Section 3.5.4;
- recommendations for selecting the eligible child in households with more than one eligible child in Section 3.5.5; and
- options available for determining the number of clusters and children per cluster are illustrated in Annex C.

The revised manual provides guidance on undertaking an immunization cluster survey covering:

- planning necessary for conducting the survey ("long range") and planning of the survey itself ("short range");
- fieldwork;
- data management and analysis;
- interpretation of findings;
- annexes covering:
  - sample size determination (Annex C);
  - budgeting for the survey (Annex F);
  - identification of survey clusters (Annex D);
  - sample forms and data collection tools used in the survey including guidance on how to complete them (Annex G, Annex H, Annex I, Annex J, Annex K); evaluating immunization coverage survey data and the indicators that can be extracted from "standard" coverage survey variables (Annex L).

Annex A provides a synopsis of basic principles of conventional sampling strategies, and is not required reading in order to prepare to undertake a survey; it is background material only.

1.3 Use of coverage surveys

A coverage survey is useful where there are no routine immunization reports in an area or to confirm the reliability of the existing reports. In addition, it provides the opportunity for the health workers to come into contact with people who are not receiving immunization – and possibly other health services – and to begin to find out why.

The most important reason for an immunization survey is to provide information on the delivery and impact of immunization services. Assessments should be planned so that they yield information targeting population groups and delivery services where corrective actions are needed. Subnational surveys may be more desirable than a single national assessment. Subnational assessments have, however, to be undertaken

uniformly and in as short a time as possible, and must cover the whole country if they are to be used to obtain a national figure. Individual subnational surveys, for example a district level survey, may be done to provide information on immunization services in those specific districts.

Box 1 gives the major reasons for undertaking an immunization coverage survey.

Box 1: Reasons for undertaking immunization coverage survey

Immunization coverage surveys may be undertaken for any of the following reasons:

As supplemental information to compare with administrative coverage reports

A survey could be undertaken to assess the quality of reports on vaccinations done routinely in health facilities and outreach centres and provide additional information on the true immunization coverage achieved in the population. A coverage survey is also useful where for some reason there are no routine immunization reports in an area, or the existing reports are very unreliable or incomplete. In other words, an immunization coverage survey can provide vital information on "where we are" with the immunization of children and women of childbearing age.

Providing information for service assessment

A survey may be designed to assess the quality of immunization services, including access to different population groups, timeliness of protection and providers of immunization.

Assessing the change in coverage over time, geographic areas or population groups

If information is needed on the change in routine coverage over time a cluster survey may be designed to provide information that can be compared to previous levels of immunization coverage. For surveys designed to assess change over time it is important to ensure that key variables such as age groups of children, immunization schedule, and catchment area are comparable. It is recommended that such surveys should not be done too frequently within the same area. There should be at least 2–3 years' gap between surveys.

Assessing coverage achieved in a supplemental immunization activity (SIA)

A survey may be designed to assess the coverage achieved in a supplementary immunization activity.

Providing information on immunization coverage demanded by funding and other agencies

Information from immunization coverage surveys can help respond to information demands by funding agencies and technical agencies.

While this manual presents immunization coverage assessment by cluster survey, there are other methods that can be used to estimate the level of immunization coverage. Examples of such methods include:

- routine reporting by health facilities and service providers;
- lot quality assurance surveys to determine whether a population group or geographical area is likely, or not, to have a particular level of immunization coverage;
- more complex survey designs, such as multistage designs (described in Annex A);
- facility-based surveys whose results can only be generalized to the population using the facilities; and
- incorporation of information on immunization coverage into wider surveys such as the demographic and health surveys (DHS) and the UNICEF multiple indicator cluster surveys (MICS).

1.4 Computer software for coverage survey data analysis

WHO has developed an Excel tool to analyse data collected in immunization coverage surveys. The tool can be obtained at: *www.who.int/immunization_monitoring/resources/en*

1.5 Objectives of the manual

The purpose of this manual is to provide the necessary skills to:

- decide if a cluster survey should be done;
- plan a cluster survey;
- conduct the fieldwork;
- organize and analyse the data; and
- interpret and use the results of the survey.

The manual describes the roles and tasks necessary in planning and implementing the survey: the *core group* guiding the planning of the survey, *coordinator* overseeing the entire survey, *field supervisors* overseeing data collection activities, *interviewers* in data collection, and *data processors* in data handling, processing and analysis.

1.5.1 Core group

The core group is responsible for preparing the terms of reference for the survey and acts as a steering group to guide the coordinator during the planning of the survey. The core group is also responsible for identifying and securing the necessary funds and for ensuring that the necessary administrative approval is given. The core group may include users of immunization coverage data from the public and private sectors and international partners in immunization activities.

1.5.2 Coordinator

The coordinator has authority over all the people involved in the exercise and has direct access to the survey commissioning authorities. The coordinator is responsible for:

- overseeing the implementation of the immunization coverage survey;
- ensuring the cooperation of other relevant government agencies;
- making budget estimates prior to identification of sources of funds for the survey;
- selecting field teams;
- executing the fieldwork; and
- reporting survey results.

1.5.3 Field supervisors

Field supervisors are responsible for ensuring that:

- local health and administrative facilities in the areas to be covered by the survey are visited before the start of work in a cluster to explain the purpose of the survey;
- interviewers are fully familiar with their task;
- each member of their teams has the necessary materials for their daily activities;
- data collection is done according to the stated instructions (see also Section **3.5.6**);
- completed forms are carefully reviewed before leaving the survey area for completeness and checking for errors;
- the completed data collection forms are given to those responsible for data processing; and
- the welfare and security of the members of the team are ensured.

1.5.4 Interviewers

Interviewers are responsible for collecting the data according to the instructions given in the data collection forms. They work under the supervision and guidance of the field supervisors. Interviewers are accountable for the data they collect and the way they collect them.

1.5.5 Data processors

The data processing team consists of people with:

- statistical expertise to help with the analysis and interpretation of the statistical results; and
- computer and data management expertise to handle the computer hardware and software including data entry.

1.5.6 Drivers

Drivers play an important role in enabling the field supervisors and the interviewers to move from one cluster to another and transport the field team to and from the field.

1.5.7 Local guides

Local guides are responsible for:

- helping field teams familiarize themselves with the clusters they are to survey;
- introducing them to the cluster administrative and social authorities;
- advising on when it is best to visit households; and
- introducing them at houses if the interviewers request it.

Their roles are NOT:

- to decide which houses are to be visited;
- to be involved in interviewing and data collection.

The table of contents shows in square brackets **[ ]** the parts of the manual that are relevant to each category of participants.

1.6 Immunization schedules

An assessment of routine immunization coverage must be based on the national immunization schedule. The coordinator, field supervisor, interviewers and data processors must be familiar with the national immunization schedule and how vaccinations are recorded on the immunization record.

Assessments of coverage achieved through supplementary immunization activities (SIA) conducted for polio eradication (e.g. national immunization days or NIDs), maternal and neonatal tetanus elimination, or measles mortality reduction may also be done. Annex E provides guidance on how the routine cluster survey methodology could be adapted for SIA assessments.

1.7 Immunization coverage survey steps

The following table shows the steps in an immunization coverage survey, the people accountable for each step and the people who undertake the work involved. The table also gives sections of the manual that users should concentrate upon according to their responsibilities in the survey.

Table 1: Summary of steps in an immunization coverage survey

Step	Responsible	Executor	Manual Section
Specifying the survey purpose	Ministry of Health	Director of public health services (or equivalent)	2.1
Survey planning	Core group/ coordinator	Coordinator	2.2–2.9
Organization of field teams and their training	Coordinator	Coordinator	3.1–3.4
Conducting the field work	Field supervisors	Interviewers	3.5–3.6
Data processing	Coordinator	Data processors	3.6
Data tabulation	Coordinator	Data processors	3.7
Data analysis	Coordinator	Coordinator	3.7
Interpreting and presenting the results	Coordinator	Coordinator and core group	5
Communicating the results	Coordinator	Coordinator	6
Using the results	Ministry of Health	Immunization Service Manager	7

2. Planning

2.1 Specifying the survey purpose *[survey commissioning authority]*

The responsible officer (whether in the Ministry of Health or not) requesting the immunization coverage survey *specifies* the purpose for the survey. See Box 1 in the Introduction.

The responsible officer then selects the core group and a coordinator.

2.2 Terms of reference for the immunization coverage assessment *[core group]*

The core group specifies the terms of reference for the assessment. See Box 2 below.

Box 2: Preparation of the terms of reference for the survey

The terms of reference specify the objective of the survey and the conditions under which it must be completed. The terms of reference define the objectives, the scope of the survey, the time frame for the survey exercise and the post survey activities.

a) Who formulates the terms of reference?

The core group formulates the terms of reference through discussions with the senior management that are calling for the survey.

b) What should the terms of reference contain?

The following is a list of specific questions that should be addressed in the discussions leading up to the terms of reference.

i) *Who* has asked for the survey? (Indicate the office or officer who has requested the survey and to whom the results should be delivered.)

ii) *Why* is the survey being requested? (Indicate how the findings will be used and the nature of the decisions or actions to be taken on the basis of the results.)

iii) *When* must the survey findings be available? (State a time frame for the survey.)

iv) *Where* will the survey be done? (Specify the geographic or administrative areas to be covered by the survey.)

v) *Who* will constitute the survey team, and what special skills are they to provide and how well suited are they for full-time involvement in the survey exercise?

vi) *How* will the expenses of the survey be covered? (Prepare a budget for the survey exercise.)

2.3 Selecting age group of children to be evaluated [*coordinator*]

The coordinator defines the age group (see Box 3) of the children to be covered according to the survey objectives. For most routine immunization coverage surveys WHO recommends using children:

- aged 12–23 months if final primary immunization is at 9 months of age;
- aged 18–29 months if final primary immunization is at 15 months of age;
- aged 0–11 months[1] if evaluating the tetanus toxoid-containing vaccine (Td or TT) coverage among mothers and whether children were protected against neonatal tetanus at birth. Tetanus toxoid coverage evaluation usually involves interviewing women who gave birth 0–11 months previously but might also include a selection of women of childbearing age if this group was also targeted for Td (or TT) immunization.

[1] The reason the age range of 0–11 months is used for evaluating Td (or TT) coverage is that this will give you information about the most recent immunization activities (that is, those that occurred within the past year) and the protection of the most recently born children and their mothers.

Box 3: Choice of age groups for the survey

Children aged 12–23 months are the most commonly chosen target population. The 12-month period is chosen to represent the births in a 12-month or a one-year period (an annual birth cohort). Most surveys are conducted to represent the most recent performance of the immunization system so the youngest possible children are chosen. However, the age range cannot be too young or some of the children in the survey will not yet be eligible for vaccination and the resulting survey results will be lower than they actually are. For example, if all children under 12 months of age (ages 0–11 months) are eligible then the survey might include a 2-week old child. This child would not be eligible for diphtheria–tetanus–pertussis (DTP) vaccine, measles vaccine and the first post-birth dose of oral polio vaccine (OPV) and thus DTP, measles and OPV coverage would be underestimated.

On the other hand if the age range is set too old, 24–35 months for example, these children will most likely be vaccinated two years ago and the survey will not reflect the most recent programme performance.

For surveys of supplementary immunization activity coverage, the age group surveyed depends on the group targeted by the SIA, for example 0–5 years for a polio NID or 15–45 years for a Td (or TT) campaign (see Annex E).

2.4 Specifying analyses *[coordinator]*

The coordinator, in consultation with the core group, specifies the variables to be collected by the survey and which indicators will be constructed from these variables. If computer-generated tables are to be produced, the coordinator must be satisfied that they will provide all the information needed – according to the objectives of the survey. The plan for data analysis should also include threshold values for immunization indicators to be used in deciding whether targets are being achieved. Examples of such indicators are given in Annex L.

If computer-generated tables are not to be used, blank tables, which are sometimes called "dummy tables", should be prepared **indicating** how the results would be presented. A statistician (who would be a member of the data processing and analysis team), in consultation with the coordinator, should prepare the dummy tables. Preparing these tables before data collection tools are designed helps to think through the data collection exercise, tabulation and analysis processes, and is a check on whether the central issues of the survey and the immunization coverage indicators will be adequately analysed and presented.

2.5 Determining the survey design *[coordinator]*

The survey method described in this manual uses a *cluster* sampling technique. This technique allows a small number of the target population to be sampled while providing statistically valid data that can be extrapolated to the whole population. The theories behind cluster sampling are complex. They are briefly described in Annex A of the manual, including footnoted references, for further information. The coordinator should, therefore, refer to that part of the manual when determining the survey design.

2.6 Determining sample size *[coordinator]*

The coordinator has to compute the number of clusters and the number of children per cluster needed to either estimate the immunization coverage or detect differences between immunization coverage over time or between places. This is referred to as sample size determination. Details of the procedure for determining the sample size are given in Annex C. Box 4 below gives a checklist of the information the coordinator needs to compute the sample size.

Box 4: Checklist for sample size determination

- The sample should be representative of the population to which the results will be extrapolated.
- Ensure that the sample size is adequate for the required precision of the results.
- Remember that bias is a result of selection techniques and not sample size and cannot be eliminated by increasing sample size.

For sample size determinations to estimate coverage, the following must be determined, estimated or assumed beforehand:

- anticipated level of immunization coverage;
- desired precision of the estimate;
- the level of *statistical* confidence of the estimate (*confidence level*);
- magnitude of differences of coverage among and within the clusters (*design effect*, defined in the glossary).

For sample size determinations to test for a difference in immunization coverage over time, the following must be determined, estimated or assumed beforehand:

- the "previous" immunization coverage level;
- anticipated level of immunization coverage;
- acceptable probability of incorrectly rejecting the hypothesis that there is no difference in immunization coverage over time (*significance level*, usually 0.05 or 0.01 is used);
- acceptable probability of correctly accepting a hypothesis that there is a specific difference in immunization coverage over time (*power*, usually 0.80 or 0.90 is used);
- a decision whether to look for either an improvement or deterioration in the immunization coverage over time or whether to merely look for any difference in either direction.

For sample size determinations to test for a difference in immunization coverage between places, the following must be determined, estimated or assumed beforehand:

- the difference in immunization coverage to be tested between the two places (normally this is 0, i.e. no difference);
- anticipated level of immunization coverage of the two places;
- acceptable probability of incorrectly rejecting the hypothesis that there is no difference in immunization coverage between places (significance level, usually 0.05 or 0.01 is used);
- acceptable probability of correctly accepting a hypothesis that there is a specific difference in immunization coverage between places (*power*, usually 0.80 or 0.90 is used);
- a decision whether to assume beforehand to test whether specific areas have higher coverage than other areas, or whether to merely look for a difference in coverage between places.

2.7 Identifying clusters *[coordinator]*

The coordinator defines what will constitute a "cluster" of children (or women for Td [or TT] immunization coverage surveys). A cluster is a collection of households[2] with identifiable geographical boundaries. Normally the clusters used for immunization coverage assessment are villages, town blocks, census tracts, etc. The details of how to select clusters are given in Annex D. Box 5 lists key points to be observed when selecting clusters.

Box 5: Points to be observed when selecting clusters for the survey

- The selection procedures should be simple rather than complex.
- The most recent census or the most up-to-date population data available should be used.

2.8 Budgeting for the survey *[coordinator]*

The *coordinator* has to prepare a *budget* for the survey including the training of the field teams and preparation of cluster household lists, etc. A key factor in the budget is the number of field supervisors and interviewers needed to collect the data from all the clusters within a reasonably short time. Box 6 gives general guidelines on determining the number of supervisors and interviewers and the length of the survey. Annex F gives a list of items to be included in the budget.

Box 6: Guidelines to follow in determining the number of supervisors and interviewers

- Each interview team should be composed of at least two members, so that interviewers can check each other's work and make sure information is recorded accurately and completely.
- There should be a mix of female and male interviewers (see Section 3.1, Box 7).
- One team of interviewers is expected to complete one cluster each day.
- The entire survey should be completed within one month, to ensure that the data are as uniform as possible.
- Health facility staff who give immunizations in the cluster should not be part of the field team. Nevertheless, the interviewers should be familiar with the immunization programme.
- A supervisor should not be responsible for more than two interview teams.
- The number of interviewers and the duration of the survey must be determined based on needs and available resources. For example, with 12 interviewers (maximum of 6 teams and at least 3 supervisors) you could complete a survey of 30 clusters in five days.

[2] A *household* is defined as a group of persons who live and eat together. This definition groups people most likely to be similarly influenced in the way they take care of their children.

2.9 Ethical clearance for the survey *[coordinator]*

It is important that the survey be conducted in accordance with the national policies on ethics for surveys involving human subjects. The coordinator should ensure that national authorities have approved the survey. If there is a national body for ethical review of research in human subjects, clearance for the survey should also be sought from this body as well.

3. Implementation

3.1 Selecting field workers *[coordinator]*

The coordinator has to identify and select the people who would be actively involved with the data collection. They include field supervisors and interviewers (data collectors). The coordinator must select people capable of working as members of a team and qualified to undertake their respective roles as defined by the coordinator. The coordinator should establish the required profile of the field supervisor and the data collectors for the tasks they have to perform. Box 7, below, gives the roles of the field supervisors and interviewers. The other people involved with the fieldwork include drivers and local guides. (The selection of local guides is normally not the responsibility of the coordinator, although the coordinator has to make arrangements with authorities and health facilities in the areas to be covered by the survey to ensure that local guides will be assigned as needed to the field teams.)

The coordinator decides whether or not those employed by the immunization system can be used as field supervisor and/or data collectors. Persons associated with the immunization services may unwittingly influence the way respondents reply to some questions, particularly those relating to reasons for not being immunized. Persons not associated with the immunization services, however, may fail to probe for important information on immunization ages, dates and reasons for failure.

The coordinator further decides whether field workers should include independent (foreign and/or international) participants so as to enhance the confidence and objectivity in the results of the survey.

Box 7: Roles of field supervisors, interviewers and local guides

The roles of all field supervisors and interviewers, and how they relate to each other, have to be clearly defined and explained to them by the coordinator.

Field supervisors

Field supervisors are responsible for ensuring the following.

- Each member of their teams has the necessary materials for their daily activities.
- Data collection is done according to the stated instructions.
- Completed forms are carefully reviewed for errors or blanks before leaving the survey area.
- The completed data collection forms are given to those responsible for data processing.
- The welfare and security of the members of the team are ensured.

Interviewers

Interviewers are responsible for collecting the data according to the instructions given in the data collection forms.

- They work under the supervision and guidance of the field supervisor.
- Interviewers are accountable for the data they collect and the way they collect them.

Depending on the local customs it may be necessary to have the correct mix of male and female members of the field teams. In some cultures interviews can only be conducted between people of the same sex.

Local guides

The role of a local guide is:

- to help field teams familiarize themselves with the clusters they are to survey;
- to introduce them to the cluster administrative and social authorities;
- to advise on when it is best to visit households;
- to introduce them at houses as may be requested by the interviewers;
- **not** to decide which houses to be visited; and
- **not** to be involved in interviewing and data collection.

3.2 Designing data collection tools (forms) *[coordinator]*

The coordinator is responsible for designing the data collection tools (questionnaires, data recording forms, etc.). The design of the data collection tools is guided by the aims of the survey to ensure that relevant data will be collectable using the tools. Three data collection forms are usual for a cluster survey of routine immunization. These are:

- the infant immunization cluster form (Annex G, Form G.1);
- the reasons for immunization failure cluster form (see Annex G, Form G.2); and
- the tetanus toxoid immunization of women cluster form (see Annex G, Form G.3).

Each form can be divided into three Sections:

Section 1: This records general identification data such as cluster number, date, area, range of birth dates.

Section 2: This records the data on immunization of individual children (or tetanus toxoid immunization of women). These data may be from either an immunization card or a verbal history from the mother or child's caretaker (see Annex G).

Section 3: This records identification of the interviewer, field supervisor and observations on the activities in the clusters and tally of the households visited and the time taken to complete each cluster.

Sample forms are given in Annex G and instructions on completing them in Annex H (for interviewers) and Annex I (for supervisors). Some modifications will be needed for forms for an SIA coverage survey (Annex E).

Box 8, below, gives the basic principles to be followed when designing the tools.

Box 8: Principles of designing data collection tools

- Do not overload the data collection form with unnecessary (although interesting) questions.
- The data collection form should be brief. It should be adapted to reflect the specific vaccines given in the country or area where the survey is carried out.
- The form should be designed to give reliable data by avoiding ambiguity in phrasing the questions so that each question conveys the same meaning irrespective of who is asking it.
- The form should be designed to give valid data (the responses elicited by the forms should be true and accurate and provide the intended information) by using questions that are clear, straightforward and not confusing or embarrassing to the respondent.
- Each form should have instructions on how it is to be used by field supervisors and interviewers. These data collection forms and the instructions should be translated into the language to be used by the interviewers, field-tested, and translated back into the original language without reference to the original to ensure that the translation was accurate. Allow enough time for this, before the interviewers' training starts.
- Be sure to design the form so that there is enough space to write the responses clearly.
- Think through the eventual analysis and presentation of the data and results, including giving consideration to facilitating computer data entry through the design of the form.

3.3 Training supervisors and interviewers *[coordinator]*

The coordinator is responsible for preparing the supervisors and interviewers for their respective roles. The results of a survey can be severely compromised by poorly prepared data collectors and field supervisors. Time should be invested in training field supervisors and interviewers in the conduct of the specific survey under preparation. It should not be assumed that because someone has participated in a similar survey before he/she does not need any further training. Field teams should be well trained in the way ***this*** survey will be done to satisfy its objectives. The importance of the instructions given in the data collection forms and the supportive role of the field supervisor must be emphasized to the interviewers.

Training supervisors and interviewers with experience in immunization coverage surveys requires at least one and a half days, including field practice. If they are not experienced in such surveys, up to three days may be needed to train them.

Training supervisors and interviewers should cover:

- objectives of this immunization coverage survey;
- the concepts of using a cluster survey for immunization coverage and eligibility criteria;
- structure of the data collection tool and the purpose of each item included in the tool; and
- roles and responsibilities of the field team members.

Supervisors who have not participated in a survey before must have field practice in the role of interviewers before the field practice of the interviewers themselves takes place.

A field practice should be done with the interviewers and field supervisors covering:

- identification of households: the first household and subsequent households;
- identification of target individuals;
- asking questions;
- data recording; and
- interview duration.

A first field test of the questionnaires should ideally be done before the field practice of the interviewers, perhaps with the field supervisors. It will provide information on:

- whether the questions are understandably phrased;
- whether the interviewers understand the questions and the instructions;
- the time it takes to locate an eligible child or woman and how long it takes to collect the information for each child or woman; and
- whether the design of the data collection tool (in translated version, if applicable) allows for legible recording of the data, as they are collected.

Box 9, below, gives recommended duration for training supervisors and interviewers.

Box 9: Recommended training duration, including field practice

Experienced supervisors	1–2 days
Experienced interviewers	1–2 days
Inexperienced supervisors	2–3 days
Inexperienced interviewers	2–3 days

3.4 Logistics [*coordinator*]

The coordinator has to ensure that all requirements for the survey are arranged before the fieldwork starts. The success or failure of a survey may depend on logistic issues. It is therefore important to pay special attention to the issues given in Box 10. When clusters are selected for inclusion in the survey the coordinator should obtain their key identifying characteristics such as: names of health facilities, names of notable persons in each cluster, boundaries and landmarks, etc. Field teams would use this information and the maps to physically locate the clusters, helped by local guides.

Box 10: Checklist for survey fieldwork

***Transport*:** Ensure that sufficient vehicles are available to transport the field teams and move them in the field as needed. The field supervisors should also have transport to be able to link up with the various field teams. (In rural areas, there should be one vehicle for each field supervisor.) The vehicles should be mechanically fit and well serviced for the duration of the fieldwork. Allowance should be made in the budget for fuel, maintenance and unforeseen repairs.

***Accommodation*:** Field teams' accommodation should be arranged for them rather than leaving teams or individuals to find their accommodation. The accommodation should, as far as possible, be conveniently placed to allow the teams to review their day's activities with their supervisor on return from the field.

***Meals*:** Field teams should be provided with either meals or food allowances for the duration of their stay in the field. Time can be saved if daytime meals are arranged for the teams instead of leaving individuals to arrange for their meals individually. Drinking water or other refreshments to be taken at work should be arranged for the field teams.

***Security*:** The security of the field teams should be ensured and not taken for granted. Local guides can be useful in advising the teams on places to be avoided and on local etiquette.

***General protection from the environment*:** Field teams should be provided with protective materials against rain or sun not only for their welfare but also to protect the data collection tools.

***Remuneration*:** Arrangements should be made for timely payments of remuneration agreed on between the survey organizers and the field teams to avoid breakdown of morale.

***Maps*:** It is important to have large and small-scale maps showing all the areas to be covered by the survey to be used in locating the clusters.

***Survey materials*:** There are basic materials each field team should have. They include: a sufficient number of data collection forms, writing pad, clipboard, plastic watertight covers for the data collection tools, carry-bag that can be closed, ball pens *not* ink pens, and copies of the national immunization schedules.

***Background information*:** Each field team should have information about the local authorities, who must be informed of their presence, carry a written letter of authorization to show local authorities, and understand how to obtain local permission when in the area to do the survey (usually requires a courtesy visit by the supervisor to local authorities). Each field team should also know the immunization hours at the health centres serving each cluster, so they can inform the residents of the visited households when to go for immunizations, and they should have a list of key dates over the last two years (e.g. festivals) that can be used to probe for birth and immunization dates.

***Communication facilities*:** Telephonic or other means of communication between the coordinator and the field supervisors must be established for the whole duration of the data collection period.

3.5 Fieldwork *[field supervisor]*

Field supervisors are responsible for the data collection (fieldwork). Fieldwork consists of collection of data on immunization from a sample of the target population by asking questions or through observation by interviewers according to a laid-out format, with data recorded on specially designed forms. Before individuals can be asked the questions on the designed forms:

- survey clusters have to be located;
- households to be visited have to be identified; and
- the individuals to be interviewed have to be located and persuaded to participate in the survey.

Fieldwork is a process that requires active monitoring. This is the main role of the field supervisor. Each field supervisor has to monitor the daily work of interviewers under his/her supervision through spot checks on the house visits and completed data collection forms. Any problems that occur during the field activities are best corrected immediately before the teams return to base.

3.5.1 Identification of survey clusters [field supervisor]

Survey clusters identified during the planning phase (see Section 2.7) have to be physically located. The field supervisors are responsible for identifying the clusters to be surveyed. This is best done with the help of maps (see Section 3.4) and knowledgeable informants.

If there is more than one cluster in a community, it should be divided geophysically into non-overlapping areas with clear boundaries.

Before the survey starts, if a community is known to be physically inaccessible or dangerous to work in, it should not be included in the list made up for cluster selection. If a cluster is found to be physically inaccessible after the survey has started, or to no longer exist, it may be replaced by another cluster (see Annex D), unless the access problem is temporary, in which case it can possibly be done later. Any such changes must have the prior approval of the field supervisor and must be noted and included in the field activities report.

3.5.2 Contact with the survey cluster officials [field supervisor]

An advance notice of the survey should be sent to the administrative authorities of the areas where the survey will be undertaken. Field supervisors should have copies of their respective field areas.

Fieldwork in any cluster should start with introductions to the officials of the cluster. The field supervisor should have a letter of introduction from either the coordinator or a higher authority outlining the work that the team will be doing in the community, the objectives of the survey and how the results would be used towards improving the health of the community. The letter should request the officials' permission to make the necessary contacts in the community, and for their support during the survey.

3.5.3 Selecting the starting household [field supervisor and interviewer]

The first house to be visited in each cluster should be selected at random (not haphazardly) using existing listings of household names for tax or electoral purposes, official maps, local maps especially made for the survey, etc. It is vital that the starting household is selected strictly randomly (i.e. without any bias due to the decisions of the people involved in the selection of the household). The basic principle is that, whatever method used, every household in the cluster must have an equal chance of selection as the starting household.

3.5.3.1 Rural areas

The first step is to establish a list of households in the cluster to be surveyed. Census records, tax lists and voting lists are the lists most commonly available, but *any* reasonably complete listing is acceptable. If none of these sources are available an ad hoc list should be established with the help of the community authorities, if at all possible. The list established is of *all* households in the cluster ***not*** just of households with target children. Time for doing this has to be included in the survey plan and budget. The field supervisor should identify someone trusted and knowledgeable of the area to advise whether the ad hoc households list is reasonably complete or not. When a reasonably complete list of the cluster households is available, follow the following steps to identify the starting household.

1) Number the households on the list.
2) Select a random number from one to the highest numbered household on the list (inclusive). Do this by using a table of random numbers (follow the instructions in Annex M).
3) Find the household on the numbered list whose number corresponds to the random number selected.

This will be the first household to visit.

If for some reason an "ad hoc" list of households cannot be prepared, an alternative method is the following:

1) Find the approximate geographic centre of the area.
2) Choose a random direction from the centre (identify all possible directions, and select a direction using the table in Annex M).
3) Count all households from the centre of the area to the edge of the area.
4) Randomly select a number between one and the number of households counted, and this will be the first household to visit.

This method should only be used once exhaustive efforts to construct an ad hoc list of households have failed.

3.5.3.2 Urban areas

If there are subdivisions (geographical, political) of the urban area that contain approximately equal populations or which can be grouped to obtain equal population distribution, use them.

If such subdivisions exist, number each subdivision and select a random number between one and the total number of subdivisions. The selected number will indicate the subdivision in which the initial household is located. Then follow the procedure described for rural areas in Section 3.5.3.1 to select the first household.

If there are no clear subdivisions, divide the urban area into subunits of approximately equal population, for example, blocks of about 100 houses. Do this by examining a map and discussing population distribution with government and health officials in the area. Once the subdivisions are established, number each subdivision and follow the procedure described above.

3.5.4 Selecting subsequent households [field supervisor and interviewer]

The process to be followed in selecting subsequent households after choosing the first household may vary depending on the nature of the dwellings, whether they are single or multifamily, scattered or clustered homesteads.

3.5.4.1 Single-family dwellings

After visiting the first household discussed in Section 3.5.3, the second household to be visited will be the one that is *nearest* to the first. The *nearest household* is defined as the household reachable in the shortest time on foot from the household just visited. The nearest household need not be in direct line of vision or on the same side of the street or road. If there are two or more households equally near to the one just visited, select the one on the immediate right as one stands in the doorway of the house looking out.

3.5.4.2 Multifamily dwellings

In densely populated urban areas, where more than one family live in a single dwelling, a different method for selecting the first household is used. To ensure an unbiased selection of households in buildings such as apartment blocks, the following system is recommended.

First, choose one floor at random. Then number the households on the selected floor and randomly select the first household to visit. The second household to visit is the door nearest to the first. After you have visited all the households on the floor, randomly choose a direction (that is, up or down). Visit all the households on that floor needed to complete the cluster. Continue from floor to floor visiting the next nearest floor that has not been visited previously. After the whole building has been visited, go to the nearest door of the nearest building and repeat the process.

Note: if multiple families live together (i.e. share cooking and sleeping quarters), this is defined as a *single household*, and only the *youngest* eligible child of the combined families should be included in the survey.

3.5.4.3 Hilly terrain with highly scattered dwellings

Identifying the nearest household can be problematic in hilly terrain with highly scattered dwellings. If it is not feasible to map the area with the dwellings in advance, then field teams have to rely on local informers to guide them to the nearest house in the course of identifying subsequent households.

3.5.4.4 Immunization "pocketing"

Immunization "pocketing"[3] can lead to serious bias of immunization coverage estimates. This is not easy to identify, without investment of time to become familiarized with survey areas before the data collection phase of the survey starts. If it is obvious that there is pocketing of immunization (or non-immunization), then the coordinator and supervisors must make a decision in advance that instead of moving from one household to the nearest one, in clusters with pocketing every other or every third household is selected. The number of households to skip depends on the degree of pocketing. More households should be skipped if the pocketing is heavy. Field teams should not make changes in the household identification procedures without the coordinator's consent.

3.5.5 Collection of data [field supervisor and interviewer]

Interviewers use the data collection forms described in Section 3.2. Data recorded in Sections 1 and 3 of the forms are all very important to the eventual assessment of the completeness of the immunization data and adherence to the survey protocol. The name and signature of the interviewer in Section 3 of the form is important to the field supervisor, to be able to monitor the work of the individual interviewers.

Interviewers should be instructed that in households with more than one eligible child data are to be collected on the *youngest* eligible child only.

3.5.5.1 Contact with the respondents

Although questions asked during an immunization coverage survey are not of a sensitive nature, interviewers must be mindful of their intrusion in the homes of the respondents. Field supervisors and interviewers must be instructed in the proper way of approaching respondents, seeking their cooperation in the survey. (Local authorities and guides should be consulted if there is any doubt as to how to relate to the population.) Respondents must be informed of the objectives of the survey and how important their participation and contribution is to the overall aim of the exercise. The approach is governed by:

Local customs: for example, whether the first contact in a household can only be made with head of the household or anyone available.

Good manners: for example, exchange of pleasantries and greetings before embarking on the interview.

[3] "Pocketing" is the situation where immunized children are in clumps of households.

Suitable time for the household visit: for example, visiting during mealtimes or when the respondents are busy in fields (in rural areas) or late in the evening after dark would be inconvenient for the respondents. In a Td (or TT) SIA survey of 15–45 year old women, a daytime survey may miss girls at high school.

Dress: for example, some attire such as shorts or mini skirts may be offensive in some cultures, or the untidy appearance of interviewers may discourage respondents' participation.

Concern for the respondents: for example, to remember that respondents' time is valuable and that they are not under any obligation to participate in the survey.

Every household visited should be given a copy of the national immunization schedule and the immunization timetable of the local health centre. For households with immunizable children, the reason for leaving the schedule would be to remind the mother of when to have her children vaccinated. For households without a child, the explanation would be that it is for future reference or to pass it on to any family that household knows with immunizable children.

If a respondent is not available during the first visit, and there is good reason to believe that the respondent is likely to return within the day or during the interviewers' stay in the area, the household has to be revisited. A message should be left at the household indicating when the interviewers are likely to return. Local guides can be useful in advising on availability and movement of the residents in the cluster (see Section 1.5.7).

3.5.5.2 Interviewing and data recording

The cluster forms in Annex 6 list the questions to be asked at each house and provide space to record information about the children and mothers in the cluster. Refer to the forms as you read the explanation of how to fill them out. Refer to Annex H for details on how to record the data on the forms. The forms and instructions are illustrative and should be customized as appropriate to meet the needs of the survey being conducted; however, all interviewers in the same survey should use the same forms.

3.5.5.3 Supplementary data

Information on the conduct of the survey is needed at the end of the exercise for final costing of the survey and overall assessment of how it was done. Field teams should, therefore, record on a separate sheet of paper their expenses, the places visited, any problems encountered, and the time of start and end of each day's activities. This provides a field activity report.

3.5.6 Implementing a quality control system on data collection [field supervisor and interviewer]

The *first level* of control is by the interviewers themselves. Working in pairs (see Box 6, Section 2.8), they can minimize the risk of procedural errors by alternating the duties of asking questions, observing and checking the recorded responses. The *second level* of control is the field supervisor for each cluster. During the data collection exercise the supervisor should observe and check that the interviewers are collecting and recording the data accurately and completely (see Box 7,

Section 3.1) through spot checks in the course of data gathering on the work of the interviewers and more comprehensively at the end of each day's fieldwork. The field supervisor ensures that each team is gathering data from the pre-selected clusters and on children in the correct age range. For data on routine tetanus toxoid immunization, the supervisor checks to ensure that data are collected on mothers with children not older than 12 months. The forms must be checked daily by the supervisor, definitely before interviewers leave the cluster area, so that any errors in completing the forms can be corrected. The supervisor must review every form to ensure the following.

1) The correct number of clusters has been surveyed. To do this, look through the cluster forms submitted by each team to see if there are forms for the correct number of clusters. When fewer than the required number clusters have been surveyed, the missing cluster(s) will need to be identified and surveyed.

2) The correct number of children who are 12–23 months of age has been listed on each cluster form for infant immunization and each cluster form for reasons for immunization failure. Similarly, the number of women interviewed for Td (or TT) immunization should be verified. If there are fewer than the expected number of children or women, arrange to revisit the cluster to correct the problem. A restart or continuation must be used, starting from the household next to the one visited last.

3) The correct number of mothers with infants 0–11 months of age has been listed on each cluster form for Td (or TT) immunization of women.

4) There are no blank items on the forms.

Information on the demographic pattern of the country or the area covered by the survey can be used to judge whether interviewers may be skipping households they should be visiting in the course of locating target children. Interviewers should, therefore, be asked to record the number of people in each household visited, whether it contains a child in the target age range or not. The supervisor can then work out the proportion of children in the target age range in the total population of the households visited and compare it with known demographic patterns.

3.5.7 Protection of forms [field supervisor and interviewer]

It is extremely important to protect the data collection forms from rain, tear and general destruction or loss. Completed data collection forms are extremely precious. If they are damaged it may be impossible to collect the same data again and the whole exercise may be made useless. The field supervisor should, therefore, ensure that interviewers have plastic watertight covers in which to put the forms (see Box 10, Section 3.4). Interviewers should also avoid exposing the forms, blank and completed, to rain or any other material that could damage them.

Interviewers should have:

- plastic watertight covers for the data collection tools;
- carrying bag that can be closed;
- flat writing clipboard;
- pens with non-smudging ink.

3.6 Data collation *[coordinator and field supervisor]*

Collected data are useless unless and until they are organized (collated) and analysed. Immunization coverage survey data must be analysed rapidly to be most useful.

3.6.1 Computer assisted

Note that although computerized data analysis is possible and may allow both quicker and more comprehensive data analysis, it is essential to be prepared to process data manually in case of unexpected computer hardware or software problems. However data are to be handled, the coordinator must daily review all cluster forms, after the field supervisors have themselves checked them, for completeness and correctness (e.g. that all children/women entered are within the birth date range specified for the survey).

3.6.1.1 Data review

If data are to be computerized the coordinator also reviews all the completed cluster forms for consistency of coding before handing them to the data processors. The data processing team consists of people with statistical expertise to help with the analysis and interpretation of the statistical results, and with computer and data management expertise to handle the computer hardware and software, including data entry.

3.6.1.2 Data processing *[coordinator]*

Data processing is a step between raw data (as collected in the field) and data analysis. In this manual it is assumed that some computer facilities would be available to support the survey. Even if computer facilities are not available the steps outlined here can be done manually. Proper planning of how the data will be handled is important as it can save time between fieldwork and publication of the results.

3.6.1.3 Data processing facilities

All the basic data processing for an immunization cluster survey can be easily done by hand unless more variables are included in the survey than were are indicated in the sample data collection forms in this manual. Since, however, computers are now readily and widely available they should be used to speed up the work and reduce possibilities for errors. Almost any configuration of the current family of computers can handle the data gathered during the cluster survey.

An Excel workbook template has been designed to support and assist field investigators during an immunization coverage cluster survey, and it enables one to perform the following tasks:

a) adapt the workbook to a specific survey, in terms of variables included and analysis performed;

b) perform data entry and import databases (meeting specific requirements);

c) analyse data, including stratified analysis (allowing the analysis of multiple strata surveys);

d) estimate proportions with 95% Confidence Intervals of each surveyed variable;

e) present main results of the survey graphically.

This Excel workbook is available from the WHO IVB website, together with a companion guide to facilitate its use, at: *www.who.int/immunization_monitoring/resources/en*

Physical facilities must be designated for data processing. The designated place must have facilities for safekeeping of the data forms and at least short-term storage facilities.

3.6.2 Manual data processing

If the data are to be handled manually, the field supervisors have to transfer the data from the cluster data collection forms to a set of cluster and summary forms:

- the infant immunization form (Form G.1 and Form G.4);
- the reasons for immunization failure form (Form G.2 and Form G.5).
- the tetanus toxoid immunization of women form (Form G.3 and Form G.6);

3.6.2.1 Completing the cluster forms *[field supervisor]*

To complete the cluster forms (see Annex G), the field supervisor completes the following steps.

1) checks that the immunizations recorded on the infant immunization cluster form are correctly recorded and revises item 13 if needed;
2) determines which children were fully immunized before one year of age (item 14);
3) completes item 13 of the tetanus toxoid immunization of women form;
4) revises item 6 on immunization status on the reasons for immunization failure form; and
5) completes the Total columns of the three forms.

For the step-by-step procedure on completing the cluster forms the field supervisor should refer to Annex I.

3.6.2.2 Completing the summary forms *[field supervisor]*

When the data are manually analysed, the next step in data collation is to transfer the "raw" data collected by the interviewers to summary forms. There are three types of summary forms:

1) the summary form on infant immunization (Form G.4);
2) the summary form on reasons for immunization failure (Form G.5); and
3) the summary form on tetanus toxoid immunization of women (Form G.6).

For step-by-step guidance on how to complete these forms the field supervisors should see Annex J.

Note: Not all items on the cluster form need to be transferred; and the introductory information on the three forms is identical.

3.6.2.3 Completing the evaluation forms *[coordinator]*

When the data are manually analysed, after completing the summary forms the evaluation forms allow for calculation of estimates of various indicators. There are three types of evaluation form:

1) the evaluation form on infant immunization (Form G.7);
2) the evaluation form on reasons for immunization failure (Form G.8); and
3) the evaluation form on tetanus toxoid immunization of women (Form G.9).

For step-by-step guidance on how to complete these forms the **coordinator** should see Annex K.

3.7 Tabulations and analysis *[coordinator]*

The coordinator decides on the type of tabulations to be produced by the data processing team as foreseen during the planning stages of the survey (see Section 2.4). Performing frequency distributions of all the variables and simple cross-tabulations are important steps towards formal in-depth analysis. For example, if data on a variable are "lumped" in one or two values this may be an indication that not much further information may be obtainable from that variable to warrant any further analysis. Frequency distributions on all variables also reveal "outliers" (extreme data) that need to be double-checked for possible errors.

The coordinator should seek technical support from statisticians, epidemiologists, etc. when analysing the data for maximum extraction of information from the data. The analysis should, at least, provide answers to the following questions.

- Is the target age group being reached?
- Why are parents or guardians not bringing their children for immunization?
- Are women being immunized with Td (or TT)?
- Why are women not being immunized with Td (or TT)?

The analysis of a routine immunization coverage cluster survey covers evaluation of:

- infant immunization;
- tetanus toxoid immunization of women; and
- reasons for immunization failure.

Annex L gives a list of indicators that can be computed from the survey data and examples of how they should be computed.

4. Monitoring the quality of work done by field supervisors or by a contractor *[coordinator/core group]*

The completeness and quality of the survey work done, both preparatory and in the field – and the ability to answer the question "How well was the survey done?" – depend critically on the quality of work done by the field supervisors, and therefore on if and how that work can be monitored. Similarly, if a survey has been contracted out (sometimes called a "third party survey"), it must be possible to monitor and assess the work of the contractor and of the field staff used by the contractor.

5. Interpreting and presenting the results *[coordinator]*

A standard survey provides several results on coverage for each antigen surveyed. There are usually three coverage indicators presented for each antigen:

- by card only or by card plus history;
- by different ages or for different age groups; and
- by inclusion or subtraction of invalid doses.

There is no single universally recommended coverage indicator. Whichever indicator is used should be fully defined and labelled properly, so that interpretation and comparisons can be done appropriately, using the same measurements. To ensure a maximum of accurate comparisons, all coverage indicators should be presented in the final report.

As a standard, the following should be included in every coverage survey report:

- valid coverage of fully immunized children (see Glossary) at 12 months – valid coverage by 12 months of age would exclude those vaccinations given after 12 months, would be based on card, would include only those diphtheria–tetanus–pertussis (DTP) doses with a minimum of 28 days between doses and at a minimum age of 9 months (270 days of age) for measles vaccination;
- coverage for each respective antigen by 12 months (or at birth for Td [or TT]), based on card plus history;
- coverage by card plus history by 12–23 months of age.

In summary, the report should always indicate whether invalid doses were included in the coverage figures provided, whether doses counted were by card only or by card plus history, and whether coverage figures are as of 12 months of age or as of 12–23 months of age.

The plan for data analysis should also include threshold values for immunization indicators to be used in deciding whether targets are being achieved. Since the results of an immunization cluster survey are based on a sample, they have an element of uncertainty. Confidence limits of estimates of the indicators are, therefore, important in providing a range likely to include the true value with a given probability.

Interpretation and presentation of results of surveys in areas where immunization cardholding is low is particularly important, because an analysis based on "card" only will be of very limited value, and the "card + history" analysis may also have to be treated with some caution. In such circumstances it will greatly aid interpretation and presentation if it has been possible to do verification of "history" data against local clinic records in the clusters being surveyed.

If any target is not being met then information on associated factors should be used to determine the underlying causes. The survey results should be compared with administrative reports (see Box 1). If there is a huge discrepancy, then further actions will be needed to determine why (such as a data quality audit of the administrative reporting system and an assessment of the quality of the survey conducted).

Diagrams and graphs are useful for communicating information, particularly for assessing changes over time. It is difficult to discern trends easily from tables. Survey results should, therefore, be complemented by graphical presentations. Such presentations might include a bar chart with bars showing coverage for a particular antigen that contrasts valid doses, all doses by 12 months of age, and all doses by 12–23 months of age, or a graph showing the differences in coverage for various antigens.

6. Communicating the results *[commissioning authority and coordinator]*

The coordinator prepares a report of the immunization coverage survey that communicates the findings and recommendations of the action to be taken to the commissioning authority. Box 11 gives the essential features of the report.

Box 11: Essential features of an immunization coverage survey report

The report should communicate the information obtained from the survey rather than simply be a record of the activities undertaken during the survey.

Tables, diagrams and other illustrations should be accompanied by commentary, pointing out levels and trends.

Body of the report

The body of the report should be *brief*, but concise, with the following chapters:

- an executive summary;
- survey objectives and terms of reference;
- background to the survey including its planning, methods and execution;
- results with analytical comments (elaborating the meaning and limitations of the results); and
- recommendations.

Additional material

Copies of the data collection forms should be annexed to the report, together with the sampling frame/cluster list.

Report presentation

The report should be attractively prepared and presented to encourage readership.

7. Taking action – using the results *[commissioning authority]*

A coverage survey not only informs on whether immunization coverage is high or low. It also provides information on patterns and reasons for non-immunization which may suggest how immunization services could be improved. The next step is to figure out how to improve the services, and to inform others of the survey results and any changes that are made. For example:

- If immunization coverage were low, say, only 30%, it would be desirable to try to determine the reasons for such a low coverage by analysing reasons for failure to immunize and devising corrective strategies.
- If health centre staff were not following the correct immunization schedule (e.g. the first DTP immunization – DTP1 – or measles given too early or not following recommended intervals for immunizations), they might need retraining or better supervision.
- If very few children had immunization cards, there may be a need to sensitize parents/guardians on the value of retaining the immunization cards. There may be a need to investigate the supply of cards and their distribution (are health workers giving cards to parents who don't have them for their child at the time of a vaccination visit?).
- If very few infants were protected from neonatal tetanus yet the measles coverage was quite high, it would be clear that mothers were not getting immunized with Td (or TT) when they brought their infant to the health centre for vaccination (i.e. a missed opportunity to vaccinate the mother at the time of the infant vaccination visit). A solution to this problem would be to retrain health workers about Td (or TT) immunization of women, and to more closely supervise this aspect of the immunization services and/or antenatal care.

It is important to inform others of the survey results and to discuss with them any changes planned for the services. This includes health centre workers, other providers in the area, and senior health officials. Feedback should be provided within one month, and is most effective if provided through meetings or newsletters. Feedback helps make health centre staff feel that they are an important part of the immunization services, thereby increasing their motivation to work hard.

Annex A:
Basic principles of sampling for surveys

A.1 Summary of common sampling methods

A.1.1 Introduction

Surveys of immunization coverage are carried out on a small number of the eligible population, either children for the childhood immunizable diseases or women against tetanus. The few people selected for inclusion in a survey are referred to as a *sample*. It has been demonstrated that sufficient information can be obtained from a (scientifically) well-selected sample. Levels of coverage and other immunization indicators can be estimated from such a sample. These estimates have some limitations regarding reliability, precision and validity. To minimize these limitations, it is necessary that some scientifically valid sampling methodology be employed.

The following is a summary of the conventional sampling methods, their advantages and disadvantages. The reader who wants to read more about these methods can find additional details in Section A.2.

The most commonly employed sampling schemes are:

- probability:
  - simple random
  - systematic
  - stratified random
  - cluster
- non-probability.

A.1.2 Summary of sampling methods, their advantages and disadvantages

A.1.2.1 Sampling

Advantages	Disadvantages
Sampling reduces demands on resources such as finance, personnel and materials.	There is always a sampling error.
Results are obtained more quickly.	Sampling may create a feeling of discrimination within the population.
Sampling may lead to improved accuracy of collected data; a smaller sample allows more effort to be made to reduce non-sampling errors and biases due to non-response.	Sampling may be inadvisable where every unit in the population is legally required to have a record.
Precise allowance can be made for sampling error (which can be found by calculation), although not for non-sampling errors.	For rare events, small samples may not yield sufficient cases for study.

A.1.2.2 Probability and non-probability sampling

Probability sampling
i) All individuals (elements) in the population have a known chance (probability) of selection. The chance of selection need not be the same for each individual or element.
ii) The knowledge of the selection probability is in contrast to the situation for non-probability sampling techniques, such as quota, convenience and chunk sampling.
iii) There must be an identified sampling frame, whether of individual elements or clusters of elements, from which the sample is to be drawn.
iv) Results over several repeated samples will be similar.
v) Random selection of sampling units may provide practical difficulties.
vi) Statistical inferences can be drawn.
Non-probability sampling
i) May be more convenient and less expensive to execute (sometimes this is the only feasible option) – does not require the identification of a sampling frame.
ii) The unknown probability of selection means that the sampling error cannot be calculated.
iii) May have strong bias.
iv) The results cannot be generalized.
v) Requires judgement and caution when interpreting the results.

A.1.2.3 Commonly used probability sampling methods

There are four "core" sampling methods that are widely used. Each of these methods is discussed in further detail in Section A.2.

- **Simple random sampling (SRS)** is a method wherein every member of the population being studied has equal probability of being selected for the sample.

  If the population can easily be divided into subgroups (based on geography, or some other easily observable characteristic), then three other sampling methods are available.

- **Systematic sampling** involves sampling every *k*th element in a list, and estimation follows that of simple random sampling.
- **Stratified random sampling** involves taking a sample from each subgroup, and thus provides a way in which estimates can be calculated for each of the subgroups and for the total population.
- **Cluster sampling** requires that samples be taken from only a *sample* of the subgroups (cluster sampling is sometimes referred to as "two-stage sampling" because of this). In certain cases, cluster sampling involves the observation of the entirety of the sampled sub-population.

When doing cluster or stratified sampling, it is sometimes the case that instead of taking a simple random sample of clusters or of elements of a cluster or strata, other probability selection methods are used.

A.2 Basic sampling concepts

A.2.1 Conventional sampling strategies

A.2.1.1 Introduction

This section[1] contains a brief review of the basic strategies used for sampling from human populations. It is intended to serve as a point of reference for the sample size discussions found in the manual. Readers interested in a more detailed discussion are referred to books on sampling theory which are provided in the reference section.

Generally, we are interested in measuring the characteristic(s) of a target population. To do this, we will take a set of observations on only some members of the target population. The target population is often identified via a sampling frame, or comprehensive listing of all the sampling units. The observational units may differ from the sampling units. For example, you might select your sample by selecting households, but then the observational unit would be individuals who live within the selected households.

Because populations tend to be large, and resources and time available for studies limited, it is usually not possible to study each observational unit (i.e. individual) or each sampling unit (i.e. household) comprising a population. For this reason there is little choice but to select a sample from the population and then make estimates regarding the entire population. In order for such estimates to be made, it is necessary that some scientifically valid sampling methodology be employed. In the following discussion, the most commonly employed sampling schemes are briefly reviewed.

The proportion of the enumeration units which possess some characteristic, Y, is denoted P, the mean level of some characteristic, X, over all N enumeration units is denoted μ, and the variance of the N values of X is denoted σ^2. Because N may be very large or the time or budget available to carry out the survey very limited, a sample of n units, with $n \leq N$ of the original N enumeration units in the population, must be selected. From the n selected enumeration units in the sample, the population proportion, mean and variance may be estimated by p, $\overline{x}$ and s^2. If this sample is selected at random from the population, these estimates will be "unbiased", meaning that – on average – the expected value of these sample-based estimates will be equal to the "true" values of the population parameters. Formally, that is:

$$E(p) = P$$

$$E(\overline{x}) = \mu$$

$$E(s^2) = \sigma^2$$

[1] This section is adapted from Lemeshow *S* et al., *Adequacy of sample size determination* (see References).

This means that if many (e.g. k) random samples were selected from this population, and if p, $\bar{x}$ and s^2 were computed for each of these samples, the average of the k sample proportions would equal the population proportion, P, the average of the k sample means would equal μ, and the average of the k sample variances would be σ^2. Unbiasedness is a desirable statistical property since it ensures that the sample values will, on average, be correct. However, it must be stressed that an estimate computed from any one particular sample may be quite different from the true value in the population. For example, it could happen that while true immunization coverage was 80%, in an unbiased sample of 20 children the observed coverage was 90%; however, a second unbiased sample of 20 children might yield an observed coverage of 70%. Neither value observed from these samples is equal to the true population coverage value, but on *average*, they would be equal. The concept of unbiasedness relates to repeated sampling and the corresponding averaging process.

Many sampling strategies are often compared to simple random sampling as a reference method, and they are usually evaluated based on three criteria:

1) Does the sampling method lead to a less *biased* and/or more *representative* estimate?

2) Does the sampling method decrease the *costs* (both monetary and other) of the survey?

3) Does the sampling method increase or decrease the *precision* (i.e. width of the confidence interval for the estimate) of the survey?

For evaluating the last of these criteria (precision), we define the *design effect*, DEFF. The design effect is equal to the ratio of the variance for the method being evaluated to the variance for a simple random sample of the same size:

$$DEFF = \frac{Var(\text{Alternative Method})}{Var(\text{Simple Random Sampling})}$$

If the DEFF is larger than one, the method has less precision (wider confidence intervals) than simple random sampling; if the DEFF is less than one, the method has greater precision (narrower confidence intervals) than simple random sampling.

In selecting a sampling design, a careful evaluation of the tradeoffs between increased precision, improved representativeness, and lowering costs must be made.

A.2.1.2 Simple random sampling

A simple random sample is one in which each of the ${}_{N}C_{n}$ possible samples has the same chance of being selected, i.e. $1/{}_{N}C_{n}$. In order to select a simple random sample it is necessary to do the following.

a) Construct a list (or "frame") of the N enumeration units.

b) Use a random process (e.g. random number table) to generate n numbers between 1 and N that identifies the n individuals in the sample.

c) Note:

- There are ${}_{N}C_{n}$ possible samples which can be selected from this population, where

$$ {}_{N}C_{n} = \frac{N!}{n!(N-n)!} = \frac{N(N-1)(N-2).....(1)}{[n(n-1)(n-2)...1][(N-n)(N-n-1)(N-n-2)...1]} $$

  For example, if $N = 25$ and a sample of size $n = 5$ is to be selected, there are ${}_{25}C_{5} = 53130$ possible samples.
- Simple random sampling provides the probabilistic foundation of much of statistical theory. It also provides a baseline to which other methods can be compared.
- Every sample of the same size has the same chance of being selected.
- Every sampling unit in the sampling frame has the same chance of being selected.
- Random selection from the sampling frame can be done by balloting, using a table of random numbers, or employing a computer.

1) Advantages of simple random sampling

The main advantages of the simple random sampling strategy follow:

a) It is simple to conceptualize.

b) Because every unit in the population has equal chance of being included in the sample, the sample is assured of being representative and subject only to sampling error.

c) Estimates are easy to calculate.

d) Sample size calculations are easy.

2) Disadvantages of simple random sampling

The disadvantages of simple random sampling are summarized below:

a) All N enumeration units in the population must be identified and labelled prior to sampling. This process is potentially so expensive and time consuming that it becomes unrealistic to implement in practice.

b) Sampled individuals may be highly dispersed. This suggests that visiting each of the sampled individuals may be a time-consuming and expensive process.

c) Minority subgroups of interest in the population may not be present in the sample in sufficient numbers for study.

d) It may not be the most efficient sampling design.

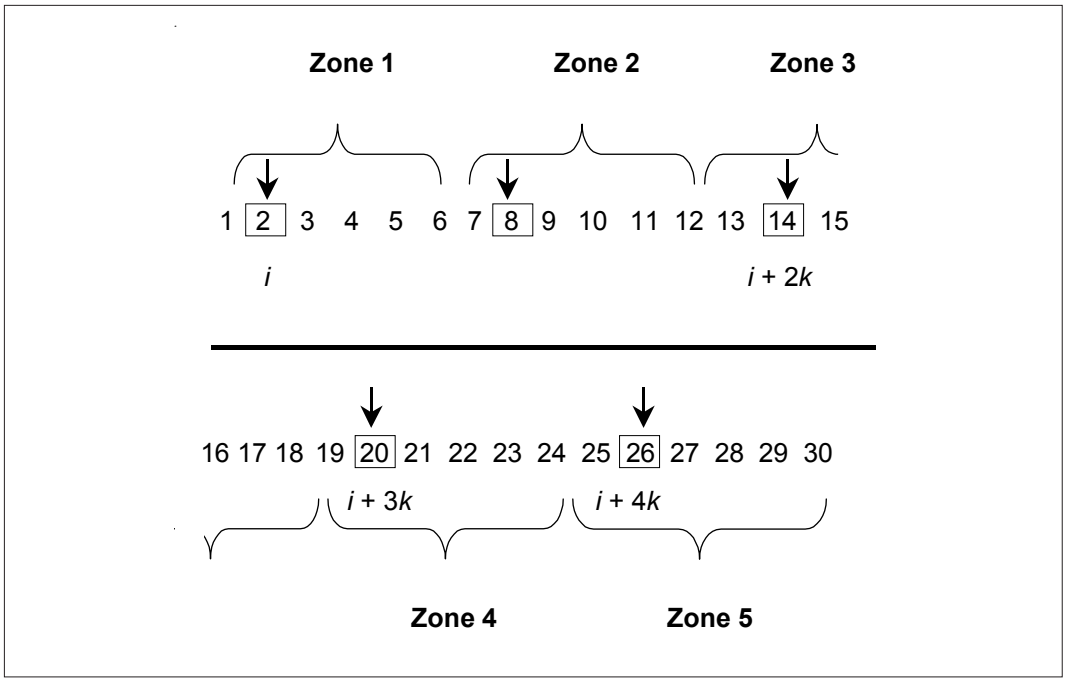
e) It ignores prior information that may be related to the variable being studied.

Alternatives to simple random sampling are often employed in actual surveys of human populations because of the major disadvantages of simple random sampling. The alternative methods may provide more precise estimates (i.e. narrower confidence intervals) for the same costs, or less precise estimates (i.e. wider confidence intervals) for substantially reduced costs.

A.2.1.3 Systematic sampling

From a list of the population or sampling frame, every *k*th unit is sampled, where 1/*k* is the sampling fraction. The first unit is selected at random from among the first *k* units. This method can save much time and effort and is more efficient in some situations than simple random sampling. For example, suppose a sample, size *n*, of records of immunizations during the past year at a local health clinic is needed for some survey of immunization. With systematic sampling, this is accomplished by creating *n* zones of $k = N/n$ records each. First, the records are arranged in some order, e.g. by date of immunization. Within the first zone, a random number *i* between 1 and *k* is selected, representing the first chosen record. Subsequent records are identified by successively adding the constant *k* to the starting random number *i*. Thus, the sample of size *n* is composed of the i^{th}, $[i+K]^{th}$, ..., $[i + (n-1)k]^{th}$ records. This is represented pictorially in Figure A-1:

Figure A-1: Pictorial presentation of systematic sampling



Systematic sampling is also a type of cluster sampling; imagine we defined the possible clusters in the population described in Figure A-1 as the sets:

$$\{1,7,13,19,25\},\{2,8,14,20,26\},\{3,9,15,21,27\},$$
$$\{4,10,16,22,28\},\{5,11,17,23,29\},\{6,12,18,24,30\}$$

Then the systematic sample corresponds to sampling the second cluster.

1) Advantages of systematic sampling

The advantage of using systematic sampling over simple random sampling may be summarized as follows:

a) The sample is easy to select.

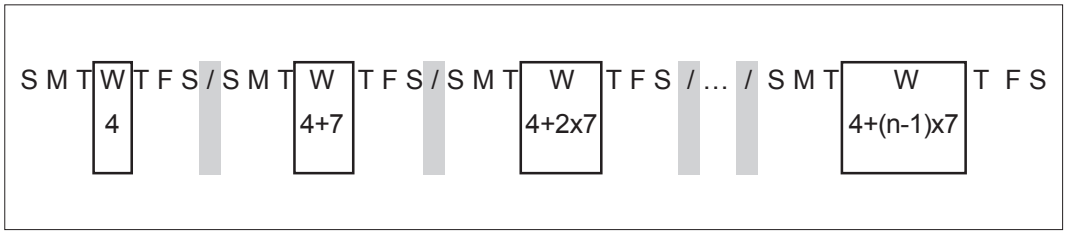
b) Using systematic sampling the selected sampling units are likely to be more uniformly spread over the whole population and may therefore be more representative than a simple random sample.

b) Under most conditions, simple random sampling formulae for parameter and variance estimates can be used with systematic sampling.

2) Disadvantages of systematic sampling

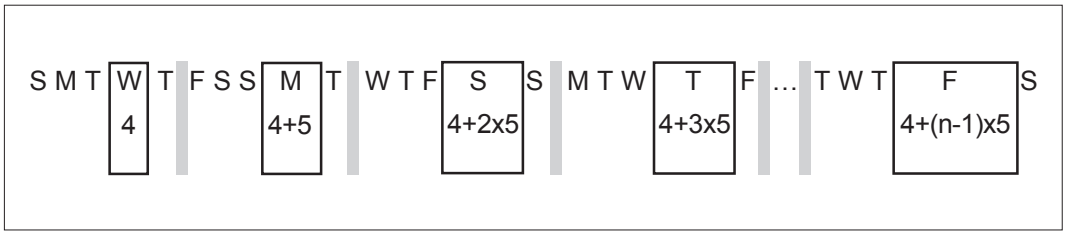
Unfortunately, there are some situations where selection of a systematic sample is ill advised. For example, if the list or frame is arranged in a cyclical fashion and *k* is the length of the cycle, a highly biased estimate will result. To illustrate, consider a study of visits to a hospital emergency room. If the emergency room has Sundays as the busiest day of the week while Wednesdays are the least busy, then the cycle is of length 7. If zones of length 7 are established, very unfortunate results may arise as seen in the Figure A-2.

Figure A-2: Systematic sampling – results from zones of length 7



Here, selecting a random number between 1 and 7 resulted in a 4, identifying Wednesday. Then repeatedly adding 7 to this random start results in the selection of successive Wednesdays, the least busy day of the week. Estimates produced from this sample will certainly not be representative of the emergency room's experience. Establishing a zone size that differs from the cycle size effectively eliminates this problem, as illustrated in Figure A-3.

Figure A-3: Systematic sampling – zone size differing from cycle size



Even when cycles do not exist, systematic sampling is often not the method of choice for actual field surveys. This is due to the fact that many of the problems listed previously with simple random sampling apply to systematic sampling as well, and it is possible to get better precision at lower cost with other methods than is possible with systematic sampling.

A brief summary of the disadvantages follows.

a) The sample may be biased if a hidden periodicity in the population coincides with that of the selection.

b) It is difficult to assess the precision of the estimate from one survey.

c) If there is a periodic trend in the data that matches the periodicity of selection, the estimate of the standard error will be too small, since the observed sample will not reflect the periodic trend.

d) If the population is ordered monotonically (i.e. in increasing or decreasing order) the estimated standard error will be too large.

A.2.1.4 Stratified random sampling

1) Definitions

A *stratum* is a subpopulation of the original population. The strata are formed on the basis of some known characteristic of the population, which is related to the variable of interest.

Some coverage surveys sample separately in urban and rural areas, for example, with each area (urban or rural) being a stratum. If a survey is being done based on health facilities, health facilities can be stratified into relatively similar groups, based on the number of beds or attending physicians at the health facility.

Stratified random sampling is the process of breaking down the population into mutually exclusive and exhaustive strata, selecting a random sample from each of the strata, and finally combining these into a single sample to estimate the population parameters. Every possible sampling unit belongs to one (and only one) stratum.

2) Sampling process

The population is first divided into groups or strata according to a characteristic of interest, for example, sex, age, geographic location, etc. A sample is then selected from each stratum using the same sampling fraction, unless otherwise prescribed for special reasons.

In order to obtain the highest precision, elements within the strata should be as homogenous as possible, while stratum-to-stratum variation should be relatively large. So strata are generally chosen by stratifying on a characteristic that relates to the quantity under study, so that the strata are internally homogeneous with respect to the quantity being studied.

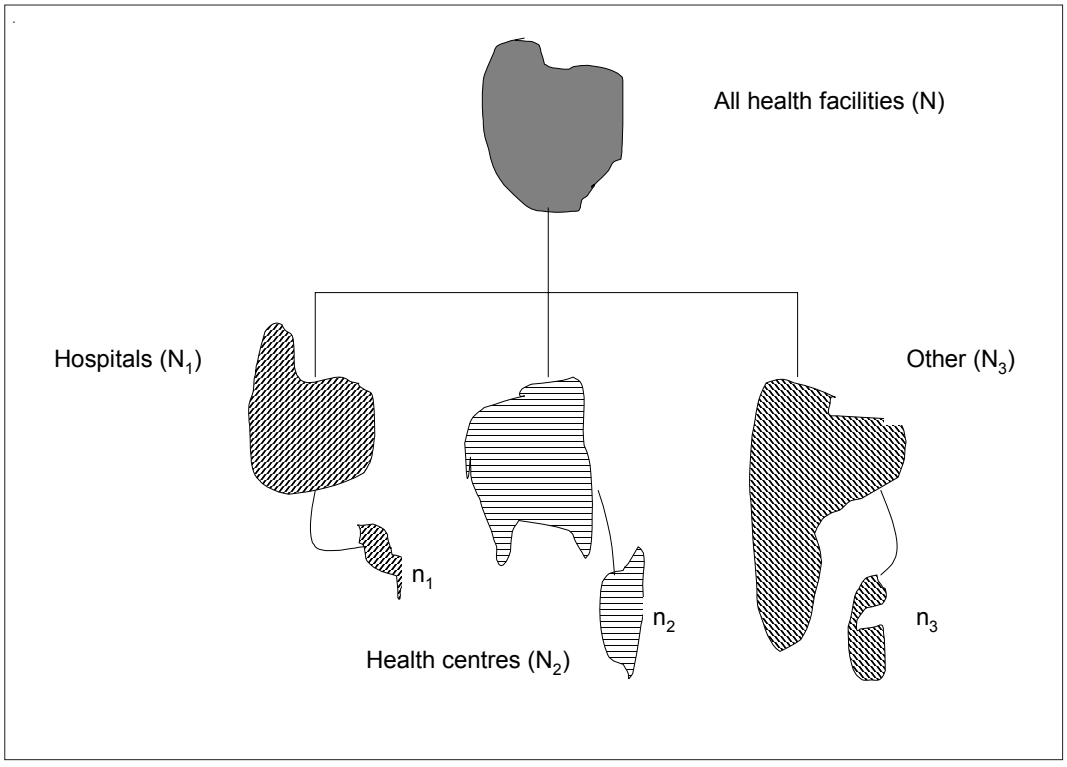
Once it is decided to use stratified random sampling, a decision must be reached as to how many elements are to be selected from each stratum. This is known as allocation of the sample. The simplest allocation scheme involves selecting an equal number of observations from each stratum. That is, $n_h=n/L$, where L is the total number of strata, and n_h is the number of elements selected from stratum h. The most commonly utilized allocation scheme is proportional allocation. In this scheme, the sampling fraction, n_h/N_h, is specified to be the same for each stratum. That is, the number of elements taken from the hth stratum is given by

$$n_h = N_h\left(\frac{n}{N}\right)$$

When proportional allocation is used, estimates of the population mean and proportion are "self-weighting". This means that when estimating the population mean, proportion or total, each sample element is multiplied by the same constant, $1/n$, irrespective of the stratum to which the element belongs.

Consider a study of all health facilities in a country. This population (of health facilities) is represented pictorially in Figure A-4 below:

Figure A-4: Pictorial presentation of stratified sampling



If a simple random sample were selected from the *N* hospitals, the large hospitals could, by chance alone, be totally missed, over sampled or under sampled. However, if a stratified sampling approach were used, then the design would involve the selection of some health facilities from each strata, assuring that each is appropriately represented in the overall sample.

3) Advantages of stratified random sampling strategy

The advantages of stratified random sampling may be summarized as follows:

- Using the same sampling fraction for all strata ensures proportionate representation in the sample of the characteristic being stratified.
- Adequate representation of minority subgroups of interest can be ensured by stratification and by varying the sampling fraction between strata as required.
- A stratified random sample may provide increased precision (i.e. narrower confidence intervals) over that which is possible with a simple random sample of the same size (i.e. DEFF<1).
- Information concerning estimates within each stratum is easily obtainable.
- For either administrative or logistic reasons, it may be easier to select a stratified sample than a simple random sample.

4) Disadvantages of stratified random sampling

- Stratified sampling requires advance knowledge of the characteristic in the population used for stratification.
- The sampling frame of the entire population has to be prepared separately for each stratum.
- Stratified sampling may not be less expensive than simple random sampling since detailed frames must be constructed for each stratum prior to sampling.
- Varying the sampling fraction between strata, to ensure selection of sufficient numbers in minority subgroups for study, affects the proportional representativeness of the subgroups in the sample as a whole, and makes the analysis of the survey more complex.
- Strata-level estimates may not have the desired level of precision, or the total sample size may be larger than needed for the population-level estimate in order to obtain the desired precision at the strata level.

A.2.1.5 Cluster sampling

Cluster sampling is a hierarchical kind of sampling in which the elementary units are at least one step removed from the original sampling clusters, and often two steps (or more) are involved.

Suppose a survey is being planned to study the prenatal care received by pregnant women in a large city.

Among the numerous problems inherent in such a study are:

- The population is very large and it might be impossible to construct an up-to-date and accurate frame. Even if one could be set up, the costs involved in setting up a detailed frame and later in attempting to contact individuals may be prohibitive.
- The population is highly dispersed. This presents significant logistical problems if there are restrictions on available time and travel expenses.

A solution to these problems is to use a cluster sampling strategy.

Sampling techniques such as simple random sampling and systematic sampling require that the sampling frames be constructed which list the individual enumeration units (or listing units).

Sometimes, however, especially in surveys of human populations, it is not feasible to compile sampling frames of all enumeration units for the entire population. On the other hand, sampling frames can often be constructed that identify groups or clusters of enumeration units without listing explicitly the individual enumeration units.

Sampling can be performed from such frames by:

- taking a sample of clusters;
- obtaining a list of enumeration units only for those clusters which have been selected in the sample; and
- selecting a sample of enumeration units.

1) Definition

Cluster sampling can be defined as any sampling plan that uses a frame consisting of clusters of sampling units. The term "cluster", when used in sample survey methodology, can be defined as any sampling unit with which one or more listing units can be associated. This unit can be geographical or temporal in nature. Typically, the population is divided into M mutually exclusive and exhaustive clusters. Unlike strata, where the strata were selected to be different from each other, but exhibit little variability within each stratum, clusters should be as homogeneous (similar) as possible, but highly variable within each cluster. Since only a subset of the clusters will be observed, each sampled cluster has to be "representative" of other non-sampled clusters, and the total variation with the population still has to be reflected in the overall estimate.

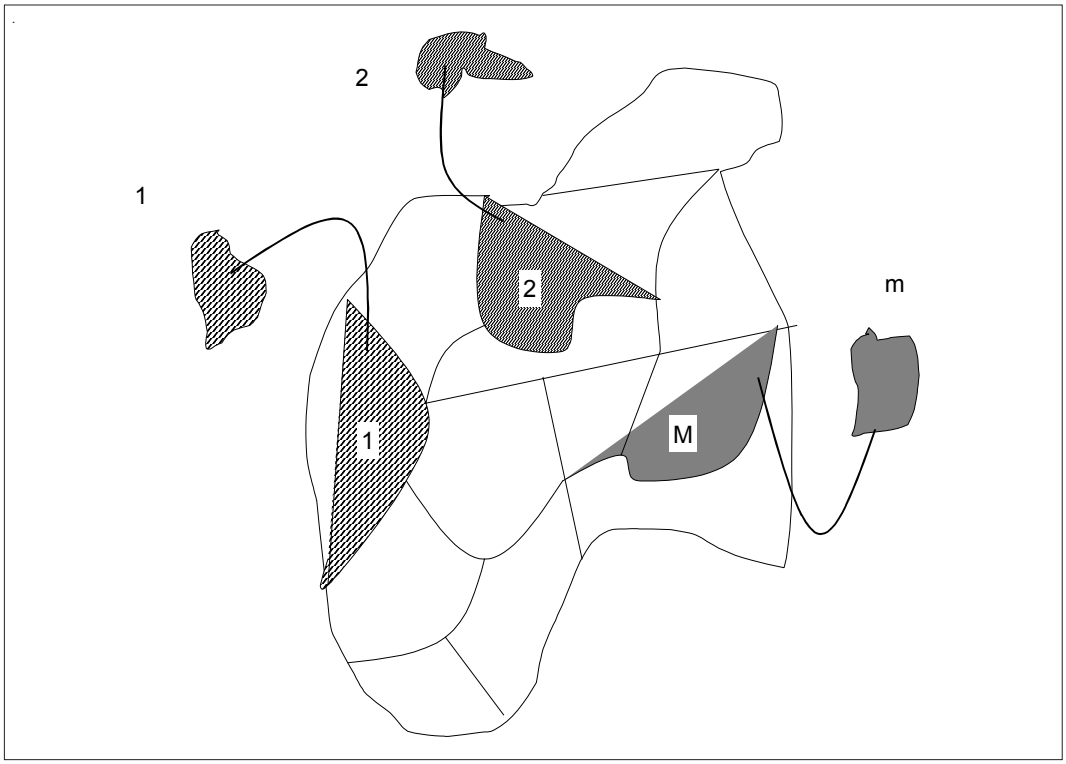
2) Sampling process

The process by which a cluster sample is selected is typically stepwise. For example, if city blocks are clusters and households are sampling units, there might be two steps involved in selecting the sample households.

- Select a sample of blocks (clusters).
- Select a sample of households (sampling units) within each block selected at the first step.

Diagrammatically, this may be represented as in Figure A-5.

Figure A-5: Schematic representation of cluster sampling



In sampling terminology, these steps are called "stages", and sampling plans are often categorized in terms of the number of stages involved. For example, a "single-stage cluster sample" is one in which the sampling is done in only one step – i.e. once a sample of clusters is selected, every sampling unit within each of the selected clusters is included in the sample. At the first stage, N_i clusters are selected from the M available clusters. Each cluster m_i has N_i sampling units; at the second stage, all N_i sampling units are studied in the ith selected cluster.

For a "two-stage sample" the m clusters are selected from the M available clusters at the first stage. At the second stage, n_i sampling units are selected, using probability sampling techniques (including simple random sampling and systematic sampling), from the ith cluster, i = 1, ..., k, ..., m. Hence samples of size n_1, n_2 , ..., m_k, ..., n_m are selected from N_1, N_2 , ..., N_k ..., N_m sampling units comprising the frames of each of the selected clusters.

Note that the total sample size (i.e. total number of sampling units included in the sample) is $n = n_1+n_2 + ... + n_k + ...+ n_m$. If $n_i = Ni$, $i=1, ...,k, ...m$, we have a "one-stage sample". On the other hand, if $n_i>N_i$ for at least some i, we have a "two-stage cluster sample".

A "multistage cluster sample" is performed in two or more steps. For example, to carry out an immunization survey of school children in a given province, the following steps might be followed:

- Select m counties from the M mutually exclusive and exhaustive counties composing the province.
- Select a sample (n_i) of townships or other minor civil divisions within each of the counties (m_i) selected at the first step.
- For each township (n_j) that was selected, select a sample (q_j) of school districts within the township.
- For each school district (q_k) that was selected, select a sample (r_k) of schools within the school district.
- Select a sample of classrooms within each of the schools selected at the fourth stage.
- Take every child within the classrooms selected at the fifth stage.

In sampling involving more than two stages, the clusters used at the first stage of sampling are generally referred to as primary sampling units or PSUs. In the preceding example, the PSU would be the counties.

With these multistage designs, writing down precise expressions for parameter estimates and associated standard errors can be difficult since each level of sampling must be accounted for. Generally, each observation must be *weighted* to reflect the probability of being included in the sample, unless probability proportional to size (PPS) sampling was implemented, in which case the sample is *self-weighting*, meaning that each sampled unit represents the same number of units in the population. In the multistage cluster example above, if at each stage the sample was taken based on simple random sampling, then the probability of a student being included in the sample would be expressed as shown below.

$\frac{m_i}{M}$	X	$\frac{n_j}{N_i}$	X	$\frac{q_k}{Q_j}$	X	$\frac{r_l}{R_k}$	X	$\frac{c_m}{C_l}$	X	$\frac{s_n}{S_n}$
Probability of selecting county		Probability of selecting township in county *i*		Probability of selecting school district in township *j*		Probability of selecting school in shool district *k*		Probability of selecting classroom in school *l*		Probability of selecting student in classroom *m*

Since all students in the classroom would be selected, $s_n = S_n$.

It should be noted that in the cluster sampling schemes described thus far, the *m* clusters were selected at random from the *M* available clusters. (Without loss of generalizability, this selection may be done systematically as well.) When clusters are selected with probability proportionate to size, denoted PPS, the determination of the probability of selection is more complex, but because the sample is self-weighting, there is no need to explicitly use the weights in the analysis. The cluster method of sampling has, however, a number of distinct advantages.

3) Advantages of cluster sampling

The advantages of cluster sampling are as follows.

a) Detailed frames need only be constructed for the *m* clusters selected at the next-to-last stage. This represents great savings in time and resources since frames need not be prepared for the entire population. Only lists of blocks or other geographic units need to be compiled, which is often feasible.

b) Cluster sampling does not usually produce as precise an estimate as simple random sampling or stratified sampling if each method were to use the same total sample size, *n*. The precision is optimized if there is intra-cluster heterogeneity and inter-cluster homogeneity (i.e. if there is little "clustering" of the parameter in question). Due to the greatly reduced cost and administrative ease, a larger cluster sample may be selected, for the same cost, than that which is possible using the other sampling schemes discussed thus far. As a result of the larger sample size, a relatively high level of precision is attained.

c) One of the most important reasons why cluster sampling is so widely used in practice, especially in sample surveys of human populations and in simple surveys covering large geographic areas, is feasibility.

d) Cluster sampling is often the most economical form of sampling.

4) Disadvantages of cluster sampling

a) There is increased complexity in the analysis of data (especially if PPS sampling was not done and there is a need to weight).

b) This relies on assumption that clusters are similar to one another (*inter-cluster homogeneity*), and that the variability across the total population is reflected in the cluster (*intra-cluster heterogeneity*).

c) If the parameter being studied is not similar across clusters, then the variance for the cluster survey will be higher than for a simple random sample (DEFF>1).

A.2.2 Complex samples

In the section on cluster sampling, a multistage design was described. In general, the building blocks of sampling design – stratified samples, cluster samples, and probability selection methods – can be combined together to build a complex sample.

Complex samples can be advisable in terms of the efficiency of design they confer – they may substantially cut down on the costs of taking the sample, and increase the resulting precision. However, estimation of means, and particularly, of sampling error, in these designs is correspondingly complex. It is recommended that you consult detailed texts on sampling should you wish to design such a sample, and work closely with a professional survey statistician on both the design and the analysis of the survey.

One specific type of complex sample that is often encountered is when you wish to aggregate the results of several surveys to obtain a combined estimate. While the help of a professional survey statistician may be recommended, a brief outline of how to do this under certain limited circumstances is given in Annex B.

A.2.3 Other types of samples

The sampling procedures described in this manual are not an exhaustive list of all the possible sampling methods that can be employed. One particular survey that has not been covered herein is the lot quality assurance survey (LQAS), which has been frequently used in health-related applications.

The goal of health workers employing LQAS procedures is to ascertain whether or not a population meets certain standards of, for example, an immunization coverage level. When lot quality assurance surveys are used to estimate overall immunization coverage they represent no more than a *stratified sampling strategy.* The reader is referred to the paper by Hoshaw-Woodard for a comparative discussion of the cluster survey versus LQAS surveys.[2]

[2] Hoshaw-Woodard S: see References.

Annex B:
Aggregating several surveys

If several surveys have been done in one global population and aggregated results are needed, you need to weight the global results by the importance of each surveyed population, or each surveyed stratum or cluster. For strata, the process is relatively simple and is explained below. For cluster surveys, this becomes more complex and it is recommended that you work with a statistician in aggregating the results.

In a "stratified" design, you've done several different surveys that cover different segments of the population – but cover the entire population.

We will start by defining terms:

$\overline{y}_i$	the estimated mean (proportion) in stratum *i*
$SE(\overline{y}_i)$	the estimated standard error of the mean (proportion) in stratum *i*
w_i	the proportion of the *total population* that is in stratum *i*
J	the total number of strata.

The formulas for calculating the estimate of a *mean* are:

$$\hat{\overline{y}} = w_1\overline{y}_1 + w_2\overline{y}_2 + \ldots + w_J\overline{y}_J = \sum_{j=1}^{J} w_j\overline{y}_j$$

$$SE(\hat{\overline{y}}) = \sqrt{\left[w_1 SE\left(\overline{y}_1\right)\right]^2 + \left[w_2 SE\left(\overline{y}_2\right)\right]^2 + \ldots + \left[w_J SE\left(\overline{y}_J\right)\right]^2} = \sqrt{\sum_{j=1}^{J}\left[w_j SE\left(\overline{y}_j\right)\right]^2}$$

Example:

You have done two EPI cluster surveys, one covering rural areas and one covering urban areas, and you have the following results.

Strata	Estimated coverage	SE of estimated coverage	% of population in strata
Rural	63%	2%	75%
Urban	74%	3%	25%

In this example, the estimated overall coverage and standard error are given by:

$$\text{Coverage} = \underbrace{(75\%)}_{w_1} \times \underbrace{(63\%)}_{\bar{y}_1} + \underbrace{(25\%)}_{w_2} \times \underbrace{(74\%)}_{\bar{y}_2} = 65.75\%$$

$$\text{SE (Coverage)} = \sqrt{\left[\underbrace{(75\%)}_{w_1} \times \underbrace{(2\%)}_{se(y_1)}\right]^2 + \left[\underbrace{(25\%)}_{w_2} \times \underbrace{(74\%)}_{se(y_2)}\right]^2} = 1.68\%$$

A 95% confidence interval for national coverage is then given by:

65.75% ±1.96 × 1.68% = (62.5%, 69.1%)

Annex C:
Determination of sample size

No prescriptive number of children is recommended in this manual. Three scenarios of sample size determination are given, for:

- estimation of immunization coverage in an area;
- testing for difference in immunization coverage over time; and
- testing for difference in immunization coverage between places.

Brief descriptions of the statistical terms such as: confidence level, significance, power, design effect, etc. are given in the Glossary.

C.1 Estimation of immunization coverage in an area

Table C-1, Table C-2, and Table C-3 give the number of children per cluster that must be sampled for desired precision of ±3%, ±5% and ±10% (respectively), for variable numbers of clusters. All of these tables assume a significance level of 5%, or confidence level of 95%, and a *design effect* of 2 (see Annex A for further description of the design effect). These are the recommended "standard" parameters for doing an immunization coverage survey, and the tables in this manual assume the default parameters. If you wish to use other values for the parameters, you will need to calculate the sample sizes needed manually.

A cluster survey is one in which the area to be surveyed is divided into sub-areas ("clusters"), and only *some* (not all) of the sub-areas are to be surveyed. See Annex A. The number of clusters, or sub-areas, which will be included in the survey, should be chosen based on criteria such as the total number of clusters in the area to be surveyed, the cost of travelling to more clusters vs sampling more children per cluster, etc.

To look up the appropriate number of children per cluster to be surveyed, you will first specify the desired level of precision of the estimates (±3%, ±5%, ±10%). This determines which table (Table C-1, Table C-2 and Table C-3, respectively)[1] you will use. You will then need to specify the expected immunization coverage (50% – 95%)[2], and the number of clusters to be surveyed, and cross-referencing these in the table will yield the number of children per cluster.

[1] Note that these tables assume a *design effect* of 2.

[2] If coverage is *less than 50%*, these tables can still be used to determine coverage; simply look up the values by subtracting the expected coverage from 1. For example, if the expected coverage is 35%, look up in the table the value for 1–35%=65%. If coverage is greater than 95%, use the sample sizes for 95%. For coverages between two values in the table (such as 73%), to be conservative, use the sample size for the *lower* of the values (in this case, 70%).

The total sample size is then the product of the number of children per cluster and the total number of clusters:

Total sample size = Number of children per cluster x number of clusters

Example: ***Estimating coverage***

If it is anticipated that the immunization coverage in the area to be surveyed is about 70%, the desired precision is ± 5% with 95% confidence, and we want to sample 40 clusters, look in Table C-2, in the row for 40 clusters and the column for 70% expected coverage; 17 children per cluster should be surveyed. The total number of children to be surveyed is 40 x 17 = 680.

At this point, you would evaluate whether or not 17 children could be found, and their parents interviewed, in one day. If 17 children is too many per cluster, it may be more feasible to look further down the column for 70% coverage; if 50 clusters are chosen, only 13 children per cluster will be needed, which may be more feasible. It is standard practice to let operational concerns – the ability to interview all the children in the cluster in one day, and the feasibility of reaching a large number of clusters – drive some of the decisions about how many clusters to use in the survey.

Table C-1: Number of children per cluster if desired precision is ±3%

Desired precision ±3%		Expected coverage									
		50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
Number of clusters	**20**	107	106	103	98	90	81	69	55	39	21
	21	102	101	98	93	86	77	66	52	37	20
	22	98	97	94	89	82	73	63	50	35	19
	23	93	92	90	85	78	70	60	48	34	18
	24	89	89	86	81	75	67	57	46	33	17
	25	86	85	82	78	72	65	55	44	31	17
	26	83	82	79	75	69	62	53	42	30	16
	27	80	79	76	72	67	60	51	41	29	16
	28	77	76	74	70	65	58	49	39	28	15
	29	74	73	71	68	62	56	48	38	27	14
	30	72	71	69	65	60	54	46	37	26	14
	31	69	69	67	63	58	52	45	36	25	14
	32	67	67	65	61	57	51	43	35	25	13
	33	65	65	63	59	55	49	42	34	24	13
	34	63	63	61	58	53	48	41	33	23	12
	35	62	61	59	56	52	46	40	32	22	12
	36	60	59	57	54	50	45	38	31	22	12
	37	58	58	56	53	49	44	37	30	21	11
	38	57	56	54	52	48	43	36	29	21	11
	39	55	55	53	50	46	42	36	28	20	11
	40	54	53	52	49	45	41	35	28	20	11
	41	53	52	50	48	44	40	34	27	19	10
	42	51	51	49	47	43	39	33	26	19	10
	43	50	50	48	46	42	38	32	26	18	10
	44	49	49	47	45	41	37	32	25	18	10
	45	48	47	46	44	40	36	31	25	18	10
	46	47	46	45	43	39	35	30	24	17	9
	47	46	45	44	42	39	35	30	24	17	9
	48	45	45	43	41	38	34	29	23	17	9
	49	44	44	42	40	37	33	28	23	16	9
	50	43	43	41	39	36	33	28	22	16	9
	51	42	42	41	39	36	32	27	22	16	8
	52	42	41	40	38	35	31	27	21	15	8
	53	41	40	39	37	34	31	26	21	15	8
	54	40	40	38	36	34	30	26	21	15	8
	55	39	39	38	36	33	30	25	20	14	8
	56	39	38	37	35	33	29	25	20	14	8
	57	38	38	36	35	32	29	24	20	14	8
	58	37	37	36	34	31	28	24	19	14	7
	59	37	36	35	33	31	28	24	19	14	7
	60	36	36	35	33	30	27	23	19	13	7
	61	36	35	34	32	30	27	23	18	13	7
	62	35	35	34	32	29	26	23	18	13	7
	63	34	34	33	31	29	26	22	18	13	7
	64	34	34	33	31	29	26	22	18	13	7
	65	33	33	32	30	28	25	22	17	12	7
	66	33	33	32	30	28	25	21	17	12	7
	67	32	32	31	30	27	24	21	17	12	7
	68	32	32	31	29	27	24	21	17	12	7
	69	31	31	30	29	26	24	20	16	12	7
	70	31	31	30	28	26	23	20	16	11	7

Table C-2: Number of children per cluster if desired precision is ±5%

Desired precision ±5%		Expected coverage									
		50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
Number of clusters	**20**	39	39	37	35	33	29	25	20	14	8
	21	37	37	36	34	31	28	24	19	14	7
	22	35	35	34	32	30	27	23	18	13	7
	23	34	34	33	31	29	26	22	18	13	7
	24	33	32	31	30	27	25	21	17	12	7
	25	31	31	30	28	26	24	20	16	12	7
	26	30	30	29	27	25	23	19	16	11	7
	27	29	29	28	26	24	22	19	15	11	7
	28	28	28	27	25	24	21	18	14	10	7
	29	27	27	26	25	23	20	17	14	10	7
	30	26	26	25	24	22	20	17	14	10	7
	31	25	25	24	23	21	19	16	13	9	7
	32	25	24	24	22	21	19	16	13	9	7
	33	24	24	23	22	20	18	15	12	9	7
	34	23	23	22	21	19	17	15	12	9	7
	35	22	22	22	20	19	17	15	12	8	7
	36	22	22	21	20	18	17	14	11	8	7
	37	21	21	20	19	18	16	14	11	8	7
	38	21	21	20	19	17	16	13	11	8	7
	39	20	20	19	18	17	15	13	11	8	7
	40	20	20	19	18	17	15	13	10	7	7
	41	19	19	18	18	16	15	12	10	7	7
	42	19	19	18	17	16	14	12	10	7	7
	43	18	18	18	17	16	14	12	10	7	7
	44	18	18	17	16	15	14	12	9	7	7
	45	18	17	17	16	15	13	11	9	7	7
	46	17	17	17	16	15	13	11	9	7	7
	47	17	17	16	15	14	13	11	9	7	7
	48	17	16	16	15	14	13	11	9	7	7
	49	16	16	16	15	14	12	11	8	7	7
	50	16	16	15	14	13	12	10	8	7	7
	51	16	15	15	14	13	12	10	8	7	7
	52	15	15	15	14	13	12	10	8	7	7
	53	15	15	14	14	13	11	10	8	7	7
	54	15	15	14	13	12	11	10	8	7	7
	55	14	14	14	13	12	11	9	8	7	7
	56	14	14	14	13	12	11	9	7	7	7
	57	14	14	13	13	12	11	9	7	7	7
	58	14	14	13	13	12	10	9	7	7	7
	59	14	13	13	12	11	10	9	7	7	7
	60	13	13	13	12	11	10	9	7	7	7
	61	13	13	13	12	11	10	9	7	7	7
	62	13	13	12	12	11	10	8	7	7	7
	63	13	13	12	12	11	10	8	7	7	7
	64	13	12	12	11	11	10	8	7	7	7
	65	12	12	12	11	10	9	8	7	7	7
	66	12	12	12	11	10	9	8	7	7	7
	67	12	12	12	11	10	9	8	7	7	7
	68	12	12	11	11	10	9	8	7	7	7
	69	12	12	11	11	10	9	8	7	7	7
	70	11	11	11	10	10	9	8	7	7	7

Table C-3: Number of children per cluster if desired precision is ±10%

Desired precision ±10%		Expected coverage									
		50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
Number of clusters	20	10	10	10	9	9	8	7	7	7	7
	21	10	10	9	9	8	7	7	7	7	7
	22	9	9	9	8	8	7	7	7	7	7
	23	9	9	9	8	8	7	7	7	7	7
	24	9	8	8	8	7	7	7	7	7	7
	25	8	8	8	8	7	7	7	7	7	7
	26	8	8	8	7	7	7	7	7	7	7
	27	8	8	7	7	7	7	7	7	7	7
	28	7	7	7	7	7	7	7	7	7	7
	29	7	7	7	7	7	7	7	7	7	7
	30	7	7	7	7	7	7	7	7	7	7
	31	7	7	7	7	7	7	7	7	7	7
	32	7	7	7	7	7	7	7	7	7	7
	33	7	7	7	7	7	7	7	7	7	7
	34	7	7	7	7	7	7	7	7	7	7
	35	7	7	7	7	7	7	7	7	7	7
	36	7	7	7	7	7	7	7	7	7	7
	37	7	7	7	7	7	7	7	7	7	7
	38	7	7	7	7	7	7	7	7	7	7
	39	7	7	7	7	7	7	7	7	7	7
	40	7	7	7	7	7	7	7	7	7	7
	41	7	7	7	7	7	7	7	7	7	7
	42	7	7	7	7	7	7	7	7	7	7
	43	7	7	7	7	7	7	7	7	7	7
	44	7	7	7	7	7	7	7	7	7	7
	45	7	7	7	7	7	7	7	7	7	7
	46	7	7	7	7	7	7	7	7	7	7
	47	7	7	7	7	7	7	7	7	7	7
	48	7	7	7	7	7	7	7	7	7	7
	49	7	7	7	7	7	7	7	7	7	7
	50	7	7	7	7	7	7	7	7	7	7
	51	7	7	7	7	7	7	7	7	7	7
	52	7	7	7	7	7	7	7	7	7	7
	53	7	7	7	7	7	7	7	7	7	7
	54	7	7	7	7	7	7	7	7	7	7
	55	7	7	7	7	7	7	7	7	7	7
	56	7	7	7	7	7	7	7	7	7	7
	57	7	7	7	7	7	7	7	7	7	7
	58	7	7	7	7	7	7	7	7	7	7
	59	7	7	7	7	7	7	7	7	7	7
	60	7	7	7	7	7	7	7	7	7	7
	61	7	7	7	7	7	7	7	7	7	7
	62	7	7	7	7	7	7	7	7	7	7
	63	7	7	7	7	7	7	7	7	7	7
	64	7	7	7	7	7	7	7	7	7	7
	65	7	7	7	7	7	7	7	7	7	7
	66	7	7	7	7	7	7	7	7	7	7
	67	7	7	7	7	7	7	7	7	7	7
	68	7	7	7	7	7	7	7	7	7	7
	69	7	7	7	7	7	7	7	7	7	7
	70	7	7	7	7	7	7	7	7	7	7

Detailed calculations

You can also estimate the total number of children to be surveyed directly. To do so, you will need to specify the *expected coverage, p*, and the desired *width of the confidence interval, d*.

To estimate the total number of children, you combine these two quantities, *p* and *d*, together with a multiplier from the normal distribution for the desired confidence level (α) in Equation 1. For $\alpha = 0.05$, $z_{1-\alpha/2} = 1.96$. You also use the design effect (the ratio between the variance from the cluster design to the variance that would be obtained from a simple random sample), so that the sample size is adequate to obtain the desired precision, given that cluster samples usually have more uncertainty than simple random samples.

Equation 1:

$$n_{\min} = DE \times \frac{z^2_{1-\alpha/2} \times p \times (1-p)}{d^2}$$

This yields the *minimum total number of children* that would have to be sampled under simple random sampling. You would then select the number of clusters *C* that you are going to sample, and determine *n* such that $n \times C \geq n_{min}$.

Example

Suppose that you wanted to know the minimum number of children to sample if the expected coverage is 80% ($p = .8$) and the desired width of the confidence interval is ±10% ($d = 10$). Then $p = .8$ and $d = 10$, and assuming a design effect of 2:

$$n_{\min} = 2 \times \frac{1.96^2 \times .8 \times (1-.8)}{10^2} = 123$$

Then if you wanted to have a sample with 30 clusters, you would need a minimum of 123/30=4.1 children per cluster – so you would round up to 5 children per cluster, for a total sample size of 30×5=150. On the other hand, if you used 20 clusters, you would need a minimum of 123/20=6.15 children per cluster, which rounded up is 7 children per cluster and a total sample size of 20×7=140.

C.2 Testing for difference in immunization coverage over time

If the objective of the survey is to test, at the 5% level of significance and a power of 90%, whether coverage in a particular area has significantly increased between two survey times then the sample size needed may be read off from Table C-4. (Details of sample size computation are in Equation 2 below.)

Table C-4: Minimum sample sizes for cluster survey to test for change in immunization coverage at 5% level of significance, 90% power (1-sided test and a design effect of 2)

P_a	P_0											
	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.40		1674	421	186	103	64	43	30	21	15	11	7
0.45	1667		1706	424	186	102	62	41	28	19	13	8
0.50	419	1704		1704	419	181	98	59	38	25	17	10
0.55	186	424	1706		1667	405	173	92	54	34	21	13
0.60	103	186	421	1674		1596	382	160	83	48	28	16
0.65	65	103	183	409	1607		1491	351	144	72	40	21
0.70	44	64	100	176	389	1506		1352	311	123	59	29
0.75	31	42	61	95	165	359	1371		1177	261	98	42
0.80	22	29	40	57	87	150	321	1201		968	203	68
0.85	16	21	27	37	52	78	131	274	996		724	134
0.90	12	15	19	24	32	44	65	107	218	756		442
0.95	8	10	12	16	20	26	35	50	79	151	478	

Where P_0 is the "old" immunization coverage

P_a is the "new" immunization coverage

It is recommended that a cluster survey covers a *minimum* of 150 children from a *minimum* of 30 clusters and a *minimum* of 5 children per cluster.

The shaded numbers in Table C-4 are less than the recommended minimum sample size of 150, which in those cases should be used instead.

Use Equation 2 to compute the sample sizes for values of P_0 and P_a not given in Table C-4.

Equation 2:

Formula for *n*

$$n = D \times \frac{\left[z_{1-\alpha/2}\sqrt{P_0\left(1-P_0\right)} + z_{1-\beta/2}\sqrt{P_a\left(1-P_a\right)} \right]^2}{\left(P_0 - P_a\right)^2}$$

with $z_{1-\alpha}$ = 1.645 (a level of 5%) and $z_{1-\beta}$ = 1.282 (β level of 20%),

and *D* is the design effect (usually 2)

Example

Estimating change in coverage: If an immunization cluster survey 24 months ago showed that the level of coverage was 60%, how many children should be sampled to test whether the level of coverage has significantly improved by at least 10% at the 5% level of significance and 90% power?

From the intersection of the column headed 0.60 and the row headed 0.70 in Table C-4, a minimum of 389 children selected in clusters would be needed. **(Note that rows and columns are not interchangeable.)**

C.3 Testing for difference in immunization coverage between places

If the objective of the survey is to test whether coverage is significantly different between two areas then the sample size needed for surveys in each area may be read off from Table C-5.

Table C-5: Minimum sample sizes per area surveyed for cluster survey to test for difference in immunization coverage between two areas at 5% level of significance, 90% power (1-sided test and a design effect of 2)

P_2	P_1																		
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.05		947	305	163	106	76	58	45	37	30	25	21	18	15	13	11	9	8	6
0.10	947		1495	433	217	134	92	68	52	42	34	28	23	19	16	14	12	10	8
0.15	305	1495		1975	545	262	157	106	77	58	45	36	29	24	20	17	14	12	9
0.20	163	433	1975		2386	639	300	177	117	83	62	48	38	31	25	20	17	14	11
0.25	106	217	545	2386		2729	716	331	192	125	88	65	50	39	31	25	20	16	13
0.30	76	134	262	639	2729		3003	776	354	202	131	91	66	50	39	31	24	19	15
0.35	58	92	157	300	716	3003		3208	819	369	209	133	92	66	50	38	29	23	18
0.40	45	68	106	177	331	776	3208		3345	845	376	211	133	91	65	48	36	28	21
0.45	37	52	77	117	192	354	819	3345		3414	853	376	209	131	88	62	45	34	25
0.50	30	42	58	83	125	202	369	845	3414		3414	845	369	202	125	83	58	42	30
0.55	25	34	45	62	88	131	209	376	853	3414		3345	819	354	192	117	77	52	37
0.60	21	28	36	48	65	91	133	211	376	845	3345		3208	776	331	177	106	68	45
0.65	18	23	29	38	50	66	92	133	209	369	819	3208		3003	716	300	157	92	58
0.70	15	19	24	31	39	50	66	91	131	202	354	776	3003		2729	639	262	134	76
0.75	13	16	20	25	31	39	50	65	88	125	192	331	716	2729		2386	545	217	106
0.80	11	14	17	20	25	31	38	48	62	83	117	177	300	639	2386		1975	433	163
0.85	9	12	14	17	20	24	29	36	45	58	77	106	157	262	545	1975		1495	305
0.90	8	10	12	14	16	19	23	28	34	42	52	68	92	134	217	433	1495		947
0.95	6	8	9	11	13	15	18	21	25	30	37	45	58	76	106	163	305	947	

Where P_1 is anticipated coverage in area 1 and P_2 is anticipated coverage in area 2.

The shaded numbers in Table C-5 are less than the recommended minimum sample size of 150, which in those cases should be used instead.

Use Equation 3, below, to compute the sample sizes for values of P_1 and P_2 not given in Table C-3.

Equation 3:

Formula for *n*

$$n = D \times \frac{\left[z_{1-\alpha}\sqrt{2P(1-P)} + z_{1-\beta}\sqrt{P_1(1-P_1)+P_2(1-P_2)} \right]^2}{\left(P_1 - P_2\right)^2} \qquad (3)$$

where

D	=	design effect, usually 2
p	=	$(P_1 + P_2)/2$
$Z_{1-\alpha}$	=	1.645 (when α = .05)
$Z_{1-ß}$	=	1.282 (when ß = .10)

Example

Estimating the difference in coverage between two places: It is desired to test whether the immunization coverage between two provinces is different and it is anticipated that the coverage in the two areas is roughly 60% and 70%. How many children should be covered by a cluster survey in each province to test for the difference at the 5% level of significance and 90% power?

From the intersection of the column headed 0.60 and the row headed 0.70 in Table C-3 a minimum of 776 children selected in clusters would be needed from each province (in a maximum 156 clusters of 5 each).

Annex D:
Identifying and selecting survey clusters

To identify clusters you must know the population of the cities, towns and villages in the area to be surveyed. The first step is, therefore, to obtain a list of all the cities, towns and villages in the area to be surveyed with as up-to-date population data as possible. Omit from the list any areas which are not going to be accessible during the survey (unless only temporarily inaccessible – see note below) or which are known not to exist any more.

This list is then the *sampling frame* from which the sample is to be selected. A sampling frame could be a list of names of villages or census tracts, or households, etc. A *sampling interval* is then determined. A sampling interval is a number used to systematically select clusters from the sampling frame. The first cluster is chosen at random (not haphazardly) using a table of random numbers (provided in Annex M with a description of how it should be used).

The steps are as follows (with reference to Table D-1):

1) Use a cluster identification form to list all communities (cities, towns, villages and sectors of cities) included in the immunization target area to be evaluated (as shown in Table D-1).

2) List the most up-to-date individual population of each community (as in column 3 of Table D-1).

3) Calculate and write in the cumulative populations as each community is added. To obtain a cumulative population, you must add the population of the next village to the combined total of all populations in preceding villages (as in column 4 in Table D-1). The final cumulative population is the same as the total population to be surveyed.

4) Calculate a sampling interval by dividing the total population to be surveyed by the number of clusters, rounding off the result to the nearest whole number. The sampling interval for the data, using an example of 30 clusters, is 4644.

   Sampling interval = Total population to be surveyed / Number of clusters

5) Select a random number which is less than or equal to the sampling interval. The number must have the same number of digits as the sampling interval. See Annex M for a table of random numbers. (The random number selected for the example from Annex M, Table M-1, is 3311.)

6) Identify the community in which cluster 1 is located. This is done by locating the first community listed in which the cumulative population equals or exceeds the random number. Write "1" beside this community in the column entitled "Cluster numbers". In the example of Table D-1 the first cluster falls within "Al Naser South" community.

7) Identify the community in which cluster 2 is located by adding the sampling interval to the random number (3311 + 4644 = 7955). The cumulative population listed for that community will equal or exceed the number you calculate. The second cluster, for the data in, also falls in "Al Naser South" community.

Cluster 2 population = Sampling interval + Random number

8) For subsequent clusters (cluster 3, cluster 4, cluster 5, etc.), identify the community in which that cluster is located by adding the sampling interval to the running total of adding the sampling interval to the random number. The cumulative population listed for that community will equal or exceed the number you calculate. So the cumulative population for cluster 3 is 7955 + 4644 = 12599, for cluster 4 is 12599 + 4644 = 17243 – both of which are in "Al Naser North" community in this example.

Cluster 3 population = Sampling interval + Cluster 2 population

Cluster 4 population = Sampling interval + Cluster 3 population

If a single community contains more than one cluster, the clusters should not overlap and should be individually identifiable. (See Section 3.5.1.)

D.1. Inaccessible clusters

If it is found during implementation of the survey, after making the sampling frame and cluster selection, that a cluster or community is not accessible:

1) If the problem is temporary (e.g. road blocked) and the place is expected to be accessible in the near future (say within a week or so), the survey there can be deferred unless there would be major problems in returning with that survey team later.

2) If the duration of the problem is not known, an alternative cluster may be selected by the field supervisor concerned, in consultation with the coordinator, in the next community listed on the sampling frame after the one containing the inaccessible cluster.

NOTE: Remoteness of a community or finding that it is going to take longer than expected to get to it is **not** a reason for counting a cluster as inaccessible. Advance planning must take account of the characteristics of the area to be surveyed, particularly if a large rural area, and provide for adequate time and logistic support to get to remote communities if they happen to be selected as clusters.

Table D-1: Example of immunization coverage survey sampling frame and cluster selection

Column 1	Column 2	Column 3	Column 4	Column 5
	Community/Area	Population	Cumulative population	Cluster numbers
1	Al Naser South	11 637	11 637	1,2
2	Al Naser North	18 181	29 808	3,4,5,6
3	Al Sinet	2 000	31 808	7
4	Hakib Alah	9 800	41 608	8,9
5	Arkaweet	4 000	45 608	10
6	Awoda	13 726	59 334	11,12,13
7	Helat Hasan	6 000	69 334	14,15
8	Al Dubasin	3 363	72 697	
9	Al Omal	12 727	85 424	16,17,18
10	Al Qatati	2 000	87 424	19
11	Al Muneera	1 500	88 924	
12	Al Mattar	2 000	90 924	
13	Al Sudani	950	91 874	20
14	Al Shartta	9 000	100 874	21,22
15	Al Muwazafin	1 500	102 374	
16	Al Zamalik	2 000	104 374	
17	Dardig	11 000	115 374	23,24,25
18	Hai Nasir	9 800	125 174	26,27
19	Al Maki	4 350	129 524	28
20	Al Gazeera	9 800	139 324	29,30
	TOTAL	139 324		
Sampling interval		$^{139,324}/_{30} = 4644$		
Random number		3311		

Annex E:
Adaptation of the routine immunization cluster survey methodology for SIA assessment

The immunization cluster survey described in the main body of the manual is designed to obtain results on immunization coverage in *routine* immunization programmes. Countries are, however, undertaking supplementary immunization activities (SIA) whose coverage needs to be assessed. Annex E summarizes the main adjustments that need to be made to the immunization cluster surveys to assess the level of coverage of SIA. The needed adaptations are presented by antigen although those needed for polio, measles and yellow fever (YF) are very similar. Table E-1 provides a comparative summary by antigen *vis-á-vis* a routine immunization coverage survey.

E.1 Polio NIDs, SNIDs, mop-ups

Supplementary immunization activity for polio typically target children aged 0–5 years old. In some mop-up campaigns, however, older age groups are also included. It is recommended, as for routine immunization surveys, that the mothers of the children in the targeted age group be interviewed. A polio SIA is typically implemented in two rounds, with about 4 to 6 weeks between the two rounds. Surveys should therefore be done after the second SIA round.

Sample size determination. High coverage levels for polio immunization are important to achieve polio eradication. Polio SIA typically reach 90% or more of the children targeted. The sample size in the survey should therefore be large enough to allow for high levels of precision in estimating the level of coverage. Table C-1, Table C-2, and Table C-3 in Annex C give guidance on the minimum sample sizes and clusters that may be needed for different anticipated levels of coverage and desired precision.

Geographic area. When planning a survey or an SIA assessment, the geographic area from which survey clusters are to be selected must correspond to the geographic area that was included in the SIA. For polio NIDs, this would be the whole country, but for SNIDs and mop-ups, only those provinces or districts where the campaign was carried out should be included in the sampling frame. (For the definition of technical terms see the Glossary.)

Age group. The age groups surveyed must correspond to the age groups targeted during the SIA. Typically these will be all children who were under the age of five at the moment of the SIA.

Children within the same family. As with the survey for routine immunization, the parent or guardian of only one child per family will be interviewed. In case several children of the targeted age group live in the same family, the selection will be done as described in the survey for routine immunization (i.e. the youngest child among the eligible children in the household).[1]

Number of doses. Polio SIA typically consist of two rounds. The survey should capture the number of doses (0, 1 or 2) that each surveyed child received in the most recent SIA.

Zero dose children. Since SIA aim at reaching previously unreached children, an important indicator of the quality of SIA is the proportion of children that had never received oral polio vaccine (OPV), the so-called "zero-dose" children. The survey should therefore include a question on whether the interviewed child had ever received polio vaccine prior to the SIA.

Use of immunization cards. Polio SIA usually do not register the doses given on an immunization card. The survey will therefore rely only on history. Special attention must be taken in designing the questionnaire so that the interviewers can differentiate between the various polio campaigns that may have taken place in the area surveyed. It is recommended that the survey should cover the most recent SIA.

Combine with routine immunization survey. If information on routine immunization is also to be recorded during an SIA assessment, the questionnaires must be designed in such a way that questions about routine immunization are asked to the subgroup of parents with children aged 12–23 months only. The sample size for this subgroup will need to be computed separately to ensure the required confidence of the estimated coverage levels. Tally sheets for the subgroup 12–23 months old must also be designed separately. Care should be taken to ensure that this procedure does not result in an overrepresentation of the children 12–23 months old in the portion of the survey on polio SIA.

Stratification. When stratification (e.g. by age, geographic area, or urban/rural status) is desired, sample sizes will need to be adjusted to ensure that coverage levels within each stratum can be estimated with the required confidence.

[1] The coordinator may decide to interview all children <5 years in a household for logistical reasons but should realize that it may also require an increase in sample size to correct for potential loss of information due to likely similarity of information on children within the same household. It is suggested that, if r% of the children in a cluster are not likely to come from households as the single eligible children in a household, the cluster sample size should be increased by a factor of $(100+r)/100$. For example if 6% of the children are likely to be multiple eligible children in households where the minimum sample size is, say, 10 then the sample should be increased to 10 x 1.06 or 11.

Output. The following two indicators should be computed:

- percentage of targeted children who were reached during the most recent polio SIA with zero dose, one dose or two doses of polio vaccine; and
- percentage of children who had never received a dose of OPV prior to the SIA and who were or were not reached by the SIA.

E.2 Measles SIA

The age group of children targeted in measles SIA is usually larger than in polio NIDs. The survey should reflect this, in that it should include all age groups targeted during the measles SIA. Measles SIA consist of only one round, and hence the survey will only measure the proportion of children immunized during the SIA (rather than number of doses received).

The methodology described for surveys after polio SIA can also be used after measles SIA.

Output. The following two indicators should be computed:

- percentage of targeted children who were immunized during the measles SIA; and
- percentage of children above the age of nine months who had never received a dose of measles vaccine prior to the SIA and who were or were not reached by the SIA.

E.3 Yellow fever SIA

Yellow fever SIA usually target large proportions of the population. As with measles, the survey should include all age groups that were part of the SIA. Only one dose of the vaccine is given in the SIA.

The same methodology as described for polio and measles can be used for yellow fever. However, as often the dose of yellow fever vaccine is marked on the vaccination card, the survey should both ask for card and for history, as for surveys on routine immunization.

Output. The following two indicators should be computed:

- percentage of the target population who were immunized during the yellow fever SIA; and
- percentage of the target population who had never received a dose of yellow fever vaccine prior to the SIA and who were or were not reached by the SIA.

E.4 Tetanus SIA

SIA with tetanus toxoid typically target all women of childbearing age (e.g. 15–44 years old) in selected districts in three rounds. The time lag between rounds is at least one month between round 1 and 2, and six months between round 2 and 3. Post Td (or TT) SIA surveys can be implemented after round 2 (to allow for the strategies to be reviewed for round 3 if needed) or after round 3 (if overall coverage results are desired).

Sample size determination. As with other surveys, the sample size is determined by the anticipated level of coverage and by the desired precision as shown in Table C-1, Table C-2 and Table C-3 of Annex C.

Geographic area. The geographic area from which clusters are to be selected must correspond to the geographic area that was included in the SIA.

Age group. The age group surveyed must correspond to the age groups targeted by the SIA. Typically these will be all women who were between 15 and 44 years old at the time of the SIA.

Women within the same family. As with the survey for routine Td (or TT) immunization, only one woman per family will be interviewed. In case several women of childbearing age live in the same family, the recommendation is to choose the youngest among them.[2]

Number of doses. Td (or TT) SIA typically consist of three rounds. The survey should capture the number of doses (0, 1, 2 or 3) that each surveyed woman received in the most recent Td (or TT) SIA, all (2 or 3) rounds combined.

Parity and "zero dose women". As SIA aim at reaching previously unreached women, an important indicator of the quality of SIA is the proportion of women who have ever been pregnant and who have never received Td (or TT), the so-called "zero-dose" women. The survey should therefore include a question on how many times the interviewed woman has been pregnant (parity) and has ever received Td (or TT) vaccine prior to the SIA (zero dose). For those women who have a child aged 0–11 months, questions should also be asked on how many doses of Td (or TT) were received during the most recent pregnancy.

Use of immunization cards. All Td (and TT) doses given, including those in SIA, are supposed to be registered on a vaccination card. Interviewers should, therefore, ask for immunization cards as with surveys for routine immunization. When cards are not available, history will be taken. Care must be taken to ensure that the questionnaire and the interviewers can differentiate between the routine immunization services and the Td (or TT) SIA that may have taken place in the area surveyed. Interviewers should therefore know the dates of the Td (or TT) SIA in the area surveyed.

[2] The coordinator may decide to interview all eligible women found in such households for logistic reasons but should realize that it may also require an increase in sample size to correct for potential loss of information due to likely similarity of information on women within the same household.

Combine with routine immunization survey. Some surveyors may wish to record information on routine Td (or TT) immunization during a Td (or TT) SIA assessment. If this is the case, the questionnaires must be designed in such a way that questions about routine immunization are asked only to the subgroup of women who have a child aged 0–11 months. The sample size for this subgroup should be computed separately to ensure ability to estimate coverage levels with the required confidence. Tally sheets for this subgroup must be designed accordingly.

Output. The following indicators should be computed:

- overall coverage with two doses of Td (or TT) by source of vaccine (either by routine immunization activities or Td [or TT] SIA);
- percentage of women who received 0, 1, 2 or 3 doses of Td (orTT) during the Td (or TT) SIA;
- percentage of zero-dose women with or without previous pregnancies who were reached by the SIA; and
- percentage of children protected against tetanus at birth for the most recent birth.

Table E-1: Recommended adaptation of routine immunization cluster survey for SIA assessment

	Routine immunization	Polio SIAs	Measles	Yellow fever	Tetanus
Desired precision	0.05–0.10	0.02–0.05	0.05	0.05	0.05
Person interviewed	Parent or guardian of the child aged 12–23 months (for childhood vaccines) or of child aged 0–11 months (for Td [or TT])	Parent or guardian of child who was in target age group at moment of SIAs	Parent or guardian of child who was in target age group at moment of SIAs	All people who were in target age group at moment of SIAs	Women who were in target age group at moment of SIAs. Additional question for mothers of child aged 0–1 months
People interviewed within household	1	1	1	1	1
Geographical area	National or as desired	Area of SIAs	Area of SIAs	Area of SIAs	Area of SIAs
Card/history	Card and history	History	History	Card and history	Card and history
Output	% coverage with BCG, OPV1, OPV3, DTP1, DTP3, measles, YF, TT2+ vaccine, % drop-out	% coverage with OPV birth dose, OPV1, OPV2, OPV3; % targeted children who never received a dose of OPV and who were vaccinated with OPV during SIAs	% measles vaccine coverage; % targeted children who never received a dose of measles vaccine prior to the SIA and who were vaccinated during SIAs	% yellow fever vaccine coverage; % targeted persons who never received a dose of yellowfever vaccine and who were vaccinated during SIAs	Overall TT2+; Protection at birth for most recent birth; TT0, TT1, TT2, (TT3) during SIAs; % women who never received a dose of Td (or TT) and who were vaccinated with a protective dose during SIAs

Annex F:
Budgeting for the survey

Note: Ensure that a contingency amount is included, in particular to cover costs of extra days that might be needed to complete work.

Budget items to be considered	Estimated cost
Preparation and production of data collection forms and instructions/guidelines	
Stationery	
Administrative support (printing, photocopies, recruiting of staff)	
Training field supervisors and interviewers Rental of training site Equipment for training (projectors, flipcharts) – rental	
Allowances for trainers	
Allowances for interviewers	
Allowances for supervisors	
Allowances for drivers and local guides	
Rental of vehicles	
Petrol for vehicles	
Communications, including costs of meetings to present results	
Consultant costs (or payment of local firm/experts to help or oversee survey)	
Total	

Annex G:
Sample immunization coverage survey forms

The number of columns provided for recording information on children or clusters is not a prescription of the number of children per cluster or the number of clusters to be surveyed. Each form should be adapted to suit the specific survey for which it is to be used.

Codes for the source of immunization, for example, should be adapted as needed, in particular if data processing is to be computerized.

Remember that if forms are to be translated this may affect the form design/ layout – direction of reading, space needed for item/row titles, and data entry.

Form G.1: Infant immunization cluster form[a]

(1) Cluster number		**(5)** Name of the child								Total	
(2) Date										Card	Card plus history
(3) Area											
(4) Range of birth dates From ..... Until											
Child number in the cluster		1	2	3	4	5	6	7	8		
(6) Birth date											
(7) Sex (M,F)											
(8) Immunization card	Yes/No										
(9) BCG	Date/+/0										
	Scar: Yes/No										
	Source										
(10) DTP1	Date/+/0										
	Source										
DTP2	Date/+/0										
	Source										
DTP3	Date/+/0										
	Source										
(11) OPV1	Date/+/0										
	Source										
OPV2	Date/+/0										
	Source										
OPV3	Date/+/0										
	Source										
(12) Measles[b]	Date/+/0										
	Source										
(13) Immunization status	Not Imm.										
	Partially										
	Fully										
(14) Fully immunized before 1 year of age (Supervisor)	Yes/No										

(15) Tally of households visited _____ **(16)** Name(s) of interviewer(s) _____

(17) Name of Field Supervisor _____

Signature: (Interviewers) _____ (Supervisor) _____

[a] Note that HepB and Hib are not listed on this form. If they are to be included in the survey, they should be added after item (11), and the same format should be used.

[b] In countries where MR or MMR are used, "measles vaccine" should be substituted accordingly.

Key:

Date/+/0:

- Date = Copy date of immunization from card, if available
- \+ = Mother reports immunization was given
- 0 = Immunization not given

Source (example)

- OUT = Outreach
- HOS = Hospital
- HC = Health centre
- PRIV = Private
- NGO = Non-governmental organization
- SIA = Supplementary immunization activity/activities

Form G.2: Reasons for immunization failure cluster form

(1)	Cluster number		**(4)** Birth dates					From			
(2)	Date							Until			
(3)	Area										
	NOTE: Ask only one question: "Why was the child not fully immunized?" Mark (**x**) the single most important reason given[a]										
	Child/woman number in cluster		1	2	3	4	5	6	7	8	Total
(5)	Sex (M or F)										
(6)	Immunization status	Not immunized									
		Partially immunized									
		Fully immunized									
(7)	Lack of information	• Unaware of need for immunization									
		• Unaware of need to return for 2nd or 3rd dose									
		• Place and/or time of immunization unknown									
		• Fear of side reactions									
		• Wrong ideas about contraindications									
		• Other									
	Lack of motivation	• Postponed until another time									
		• No faith in immunization									
		• Rumours									
		• Other									
	Obstacles	• Place of immunization too far									
		• Time of immunization inconvenient									
		• Vaccinator absent									
		• Vaccine not available									
		• Mother too busy									
		• Family problem, including illness of mother									
		• Child ill – not brought									
		• Child ill – brought but not given immunization									
		• Long waiting time									
		• Other									
(8)	Tally of households visited										
(9)	Name of Interviewer										

Signature _____

[a] If it is felt that categorization/pre-coding of possible responses may risk missing potentially important information from the respondents, the interviewers can be instructed simply to write down verbatim the reply given by the child's mother/caretaker or by the woman (provided a suitably designed format for recording the information is given to the data collectors). The survey supervisors and coordinator will later review all responses and decide on appropriate categories for presentation of the analysis.

Form G.3: Tetanus toxoid immunization of women cluster form

(1) Cluster number		**(5)** Name of the mother									Totals **(to be completed by supervisor)**
(2) Date											
(3) Area											
(4) Range of birth dates											
From											
Until											
Woman's number in the cluster			1	2	3	4	5	6	7	8	
(6) Birth date of child											
(7) Total number of lifetime pregnancies											
(8) History of Td (&/or TT) immunization in last pregnancy	(a) Number of Td (&/or TT) doses received prior to last pregnancy										Dose LAST pregnancy TT0= TT1= TT2= TT3= TT4= TT5 or more=
	(b) Number of Td (&/or TT) doses received in last pregnancy										
	(c) Card available for Td (&/or TT) received in last pregnancy? Y/N										Yes= No=
	(d) Whether or not card is available; was a card ever received? Y/N										Yes= No=
											Card / Card + history / Tally source
(9)	TT1	Date or Y or N									
		Source									OUT= HOS= HC= PRIV= NGO= SIA= SCH= WCV= OTH=
	TT2	Date or Y or N									
		Source									
	TT3	Date or Y or N									
		Source									
	TT4	Date or Y or N									
		Source									
	TT5	Date or Y or N									
		Source									
(10) Antenatal care[a]	Number of visits in last pregnancy										One visit= Two or more visits=
(11) Other visits to health facility[a]	Number of visits in last pregnancy										One visit= Two or more visits=
(12) Delivery of baby[a]	Where	Home									Home=
		Hospital/HC									HC/Hospital=
		Other									Other=
	By who	Health staff									Health staff=
		TBA									Traditional attendant=
		Other									Other=
		Nobody									Nobody=
(13) Child protected against neonatal tetanus[a] **(supervisor)**	Yes by card (Y)										Yes by card =
	Yes by card+history										Yes by C+H=

(14) Tally of households visited _____

(15) Name(s) of interviewer(s) _____ **(16)** Name of Field Supervisor _____

Signature(s) _____ Signature _____

[a] Only indicate data related to "**last** pregnancy", i.e. pregnancy that led to a child now aged 0–11 months.

For key to sources, see Form G.1

Form G.4: Summary of infant immunization[a]

(1) Date of first interview _____ **(2) Date of last interview** _____ **(3) Area** _____ **(4) Age group evaluated** _____

Cluster number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Total
(6) Number in cluster																															
(7) Card– Yes																															
(8) BCG																															
Card only																															
Card+history																															
BCG scar																															
Source: HOS																															
HC																															
OUT																															
PRIV																															
NGO																															
SIA																															
(9) DTP 1																															
Card only																															
Card+history																															
DTP 2																															
Card only																															
Card+history																															
DTP 3																															
Card only																															
Card+history																															
Source: HOS																															
HC																															
OUT																															
PRIV																															
NGO																															
SIA																															

[a] If HepB and Hib are included in the survey, add lines for recording summary information for these antigens following item 10 (OPV), using the same format.

(continued)

Form G.4: Summary of infant immunization *(continued)*

Cluster number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Total
(10) OPV1																															
Card only																															
Card+history																															
OPV2																															
Card only																															
Card+history																															
OPV3																															
Card only																															
Card+history																															
Source: HOS																															
HC																															
OUT																															
PRIV																															
NGO																															
SIA																															
(11) Measles																															
Card only																															
Card+history																															
Source: HOS																															
HC																															
OUT																															
PRIV																															
NGO																															
SIA																															
(12) Immunization history																															
Not immunized																															
Partially immunized																															
Fully immunized																															
(13) Fully immunized before age 1																															
(14) Households visited																															

For key to sources, see Form G.1

Form G.5: Summary of reasons for immunization failure

(1) Date of first interview _____ (2) Date of last interview _____ (3) Area _____ (4) Age group evaluated _____

Cluster number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Total
(5) Immunization status																															
Not immunized																															
Partially immunized																															
Fully immunized																															
(6) Lack of information																															
a) Unaware of need for immunization																															
b) Unaware of need to return for 2nd / 3rd dose																															
c) Place and/or time of immunization unknown																															
d) Fear of side reactions																															
e) Wrong ideas about contraindications																															
f) Other																															
Subtotal																															
(7) Lack of motivation																															
g) Postponed until another time																															
h) No faith in immunization																															
i) Rumours																															
j) Other																															
Subtotal																															

(continued)

Form G.5: Summary of reasons for immunization failure *(continued)*

Cluster number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Total
(8) Obstacles																															
k) Place of immunization too far																															
l) Time of immunization inconvenient																															
m) Vaccinator absent																															
n) Vaccine not available																															
o) Mother too busy																															
p) Family problem, including illness of mother																															
q) Child ill – not brought																															
r) Child ill – brought but not given immunization																															
s) Long waiting time																															
t) Other																															
Subtotal																															
TOTAL																															

Form G.6: Summary of tetanus toxoid immunization of women

(1) Date of first interview _____ **(2) Date of last interview** _____ **(3) Area** _____ **(4) Age group evaluated** _____

Cluster number		**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**Total**
(5) Number in cluster																																
(8) Td (or TT) last pregnancy	TT0																															
	TT1																															
	TT2																															
	TT3																															
	TT4																															
	TT5 or >																															
(8) Card available																																
(8) Card ever received																																
(9) Life time Td (or TT)	TT1 card																															
	TT1 card + history																															
	TT2 card																															
	TT2 card + history																															
	TT3 card																															
	TT3 card + history																															
	TT4 card																															
	TT4 card + history																															
	TT5 card																															
	TT5 card + history																															
(9) Source of Td (or TT)	OUT																															
	HOS																															
	HC																															
	PRV																															
	NGO																															
	SIA																															
	SCH																															
	WCV																															
	OTH																															

(continued)

Form G.6: Summary of tetanus toxoid immunization of women *(continued)*

Cluster number		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Total
(10) ANC	1 visit																															
	2 or more visits																															
(11) Other visits	1 visit																															
	2 or more visits																															
(12) Delivery location	Home																															
	Hospital/HC																															
	Other																															
(12) Delivery assistance	Health staff																															
	TBA																															
	Other																															
	Nobody																															
(13) Protected at birth	By card																															
	By card+history																															
(14) Households visited																																

Key:

ANC = antenatal care

For key to sources, see Form G.1

Form G.7: Evaluation of infant immunization[a]

(1) Area		**(2)** Age group evaluated	
(3) Date of first interview		**(4)** Date of last interview	
(5) Number in survey			

	Total card		Total card plus history	
	Number	**Percentage**	**Number**	**Percentage**
(6) BCG				
BCG scar				
Source: HOS				
HC				
OUT				
PRIV				
NGO				
SIA				
(7) DTP 1				
DTP 2				
DTP 3				
Source: HOS				
HC				
OUT				
PRIV				
NGO				
SIA				
(8) OPV1				
OPV2				
OPV3				
Source: HOS				
HC				
OUT				
PRIV				
NGO				
SIA				
(9) Measles[b]				
Source: HOS				
HC				
OUT				
PRIV				
NGO				
SIA				
(10) Immunization history:				
Not immunized				
Partially immunized				
Fully immunized				
(11) Fully immunized before one year of age				
(12) Total number of households				
(13) Average number of households per cluster				

[a] If HepB or Hib are included in the survey, add lines to this form after the lines for OPV, using the same format.

[b] In countries where MR or MMR are used, "measles vaccine" should be substituted accordingly.

For key to sources, see Form G.1

Form G.8: Evaluation of reasons for immunization failure

(1) Area		**(2)** Age group evaluated	
(3) Date of first interview		**(4)** Date of last interview	
(5) Number in survey			

	Total	Percentage
(6) Partially or not immunized		
(7) Lack of information		
a) Unaware of need for immunization		
b) Unaware of need to return for 2nd or 3rd dose		
c) Place and/or time of immunization unknown		
d) Fear of side reactions		
e) Wrong ideas about contraindications		
f) Other		
Subtotal		
(8) Lack of motivation		
g) Postponed until another time		
h) No faith in immunization		
i) Rumours		
j) Other		
Subtotal		
(9) Obstacles		
k) Place of immunization too far		
l) Time of immunization inconvenient		
m) Vaccinator absent		
n) Vaccine not available		
o) Mother too busy		
p) Family problem, including illness of mother		
q) Child ill – not brought		
r) Child ill – brought but not given immunization		
s) Long waiting time		
t) Other		
Subtotal		

Form G.9: Evaluation of tetanus toxoid immunization of women

(1) Area		**(2)** Age group of children whose mothers are to be evaluated	
(3) Date of first interview		**(4)** Date of last interview	
(5) Number in survey			

		Total card		Total card plus history	
		Number	**Percentage**	**Number**	**Percentage**
(8) Td (or TT) last pregnancy	TT0				
	TT1				TT2+:
	TT2				
	TT3				
	TT4				
	TT5>				
(8) Card available					
(8) Card ever received					
(9) Life-time Td (or TT)	TT1				
	TT2				
	TT3				
	TT4				
	TT5>				
(9) Source Td (or TT)	OUT				
	HOS				
	HC				
	PRIV				
	NGO				
	SIA				
	SCH				
	WCV				
	OTH				
(10) ANC	1 visit				
	2 or more visits				
(11) Other visits	1 visit				
	2 or more visits				
(12) Delivery location	Home				
	Hospital / HC				
	Other				
(12) Delivery assistance	Health staff				
	TBA				
	Other				
	Nobody				
(13) Protected at birth					
(14) Total number of households					
(15) Average number of households per cluster					

Key for sources, see G-1

Annex H:
Interviewer instructions for completing the cluster forms

NOTE: These instructions must be revised according to design of the actual forms to be used, and be available for the training of survey personnel.

The cluster forms list the questions to be asked at each house and provide space to record information about the children and mothers in the cluster.

Refer to the forms in Annex G as you read the explanation of how to fill them out. The first four items on the three forms, for recording general identification data, are identical. These items are completed as follows:

Item 1	Record the cluster number.
Item 2	Record the date of the interview.
Item 3	Identify the city, town or village of the cluster by referring to the cluster identification form.
Item 4	Identify the birth date range of children who will be evaluated in the survey. These dates will be based on the date of the interview. Two different ranges of birth dates are computed as follows.

H.1 Birth date range cluster forms related to infant immunization (Forms G.1, G.2)

To determine the *earliest* acceptable birth date, subtract exactly 24 months from the date of the survey. (Twenty-four months instead of 23 months are subtracted because all children who are even one day less than 24 months of age have to be included. By subtracting 24 months, children who are exactly 24 months of age will also be included. This is an acceptable error.)

To determine the latest acceptable birth date, subtract exactly 12 months from the date of the survey.

The following example illustrates how this is done.

- Assume a survey date of 15 May 2000.
- Count back from the survey date exactly 24 months to determine the earliest acceptable birth date.
- Count back from the survey date exactly 12 months to determine the latest acceptable date.

Example: To determine the latest acceptable birth date range

Months	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M
No. of months	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
	Earliest acceptable d.o.b: 15 May 1998												Latest acceptable d.o.b: 15 May1999												Interview: 15 May 2000

The shaded area in the figure above represents the birth dates of the children to be evaluated if the interview date is 15 May 2000 (that is, birth dates falling on or between 15 May 1998 and 15 May 1999). **Note:** If no immunization cards or birth records are available, use months of birth instead of specific dates. (Local festival calendars can help fix dates of birth.)

H.2 Birth date range for the tetanus toxoid immunization of women cluster form (Form G.3)

To determine the earliest acceptable birth date, subtract exactly 12 months from the date of the interview. The latest acceptable birth date is the date of the interview.

After completing items 1–4 on all three forms, visit the first household, and proceed as follows.

i) Introduce yourself and ask to see the head of the household. If the head of the household is not present, ask to speak to the spouse, another adult or a mature child.

ii) Explain what you are doing and why you will be asking them questions.

iii) Questions of immunization should, ideally, be directed to the mother of the child. If the mother is not available then the question may be addressed to an adult in the household most likely to be able to provide the data.

iv) Ask the ages of the children living in the household.

v) Determine if there are any resident children in the household aged 0–23 months. (A resident child is defined as one who spent the previous six months in the household, although the child does not have to be present for the interview.) Then:

vi) Use the chart below to decide which cluster form to use.

If there is a child who is:	then:
0–11 months	Complete the cluster form for tetanus toxoid immunization of women and reasons for immunization failure.
12–23 months	Complete the cluster forms for infant immunization and reasons for immunization failure.
24 months or more	Do not fill out any form (unless the survey is covering, for example, 0–5 year olds for polio NIDs, for which separate forms or questions will be needed). Tally the household visited on the cluster forms for infant immunization and tetanus toxoid immunization of women. Go to the next nearest house and begin again.

Specific guidelines on completing each of the three forms follow. The numbers listed in the guidelines correspond to the item numbers on the cluster forms.

H.3 The cluster form for infant immunization (Form G.1)

Item 5 Write the name of the youngest child in the household in the age range of 12 to 23 months. There are eight columns for recording the data in the sample form shown in Annex G.[1] If more than eight children are to be recorded a second form should be used as a continuation. The data collection form should be designed according to the number of children to be sampled as indicated in Section 2.6, with at least one additional column. If, in a household with the child who completes the desired number of children to be sampled, there is another child who is 12–23 months, then list that child here in the additional column. Also it is useful to take an "extra" eligible child anyway, especially on the first day of a survey, in case an ineligible child is taken by mistake.

Item 7 Write "M" if the child is male, "F" if female.

Item 6 Write the birth date for each child listed.

Item 8 Ask to see the immunization card(s) for each child on the list. (It is possible that a single individual may have several immunization cards). Write "Yes" in the appropriate box if the child has an immunization card. Write "No" if the child does not have an immunization card.

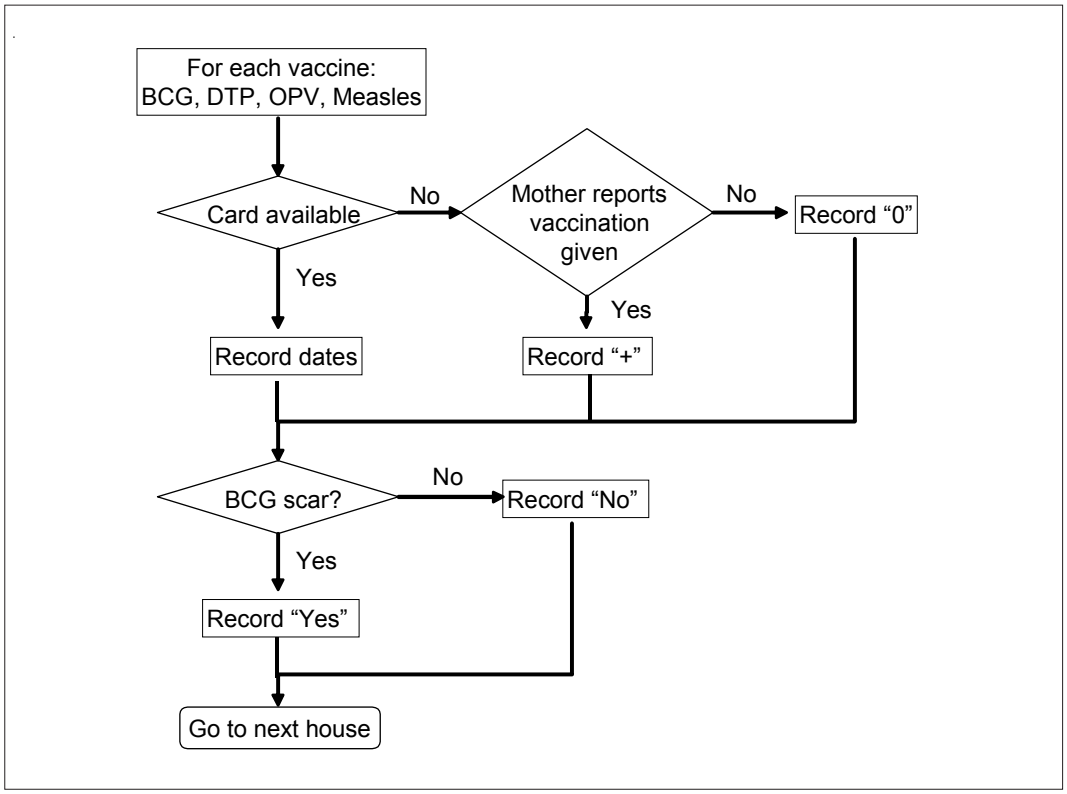
Items 9–12 The row "DATE/+/0" is completed by using a coding system. If the immunization card is present and the immunization was given, copy the date of the immunization. This is known as verifying the immunization "by card". If the immunization card is not present, ask the mother whether the child received the immunization. If the mother reports that the immunization was given, write "+" in the box. This is known as verifying the immunization "by history". If the immunization was not given, write "0" in the box. It does not matter whether that information is obtained from the mother or from the immunization card. If the immunization card is present, you may ask the mother or care taker whether the child who received immunizations that were not recorded on the vaccination card (e.g. doses given during supplementary immunization activities). In such case, write "+" in the respective box or boxes.

In some circumstances it may be possible or desirable to try to verify "by history" information from immunization/health records held at a health centre or clinic that serves the locality being surveyed. This has the advantage of increasing the number of children/women for whom immunization dates are available, and therefore improves the quality of the analysis that can be done and conclusions. (In this case item 8 needs modification.) It is particularly useful to make time to do this in areas where it is known or found early that immunization cardholding at home is very low, but that there are accessible clinic records for the period (up to two years ago).

See Figure H-1 below for a flow chart for completing items 9–12.

[1] This form, and all the forms in the manual, should be redesigned to accommodate the peculiar requirements of ***this*** survey such as space for the names and comments.

Figure H-1: Flow chart for completing interview questions



Ask the mother where the child received the immunization, and record it in the relevant box in the row titled "Source". For example, use the following abbreviations (suggested list but can be modified):

HOS	hospital
OUT	outreach
HC	health centre
PRIV	private
NGO	non-governmental organization
SIA	dose received during supplementary immunization activities (For doses received during SIA, only mark "SIA" and not "HOS", "OUT", "HC", "PRIV" or "NGO").

If the mother does not remember where the child received the immunization, mark "0" in the box.

After listing the information on all the children in the household aged 12–23 months, check the data recorded for any obvious errors. For example, check for:

- blank spaces, which would indicate questions either not asked or data simply not recorded;
- immunization dates that occurred prior to the date of the child's birth; and
- duplicate recordings, which would indicate that the same data were mistakenly recorded for more than one child.

Item 13 Determine whether the child is fully immunized or not. Count immunizations that were recorded by card and by history. The immunization status of the child is recorded as **fully, partially** or **not** immunized by marking "X" in the relevant box.

Mark	if the child received
Fully immunized [a]	3 DTP, 3 polio, 1 BCG, and 1 measles plus 3 hepatitis B and 3 Hib if these vaccines have been part of the national or local immunization schecule for at least two years
Partially immunized	At least one immunization
Not immunized	None of the immunizations

[a] If one wants to assess coverage of an antigen not yet part of the immunization programme for two years, modify the acceptable age and birth date range to be evaluated to allow for "up-to-date immunization for age" rather than "fully immunized".

Item 14 Interviewers will leave this item blank; the supervisor will complete it.

Item 15 Tally each household you visit. This keeps track of the total number of houses visited to find seven (or the specified number of) children and/or women in the cluster.

At this point, before the interview ends, complete the cluster form for reasons for immunization failure. You must complete item 5 on this form for each of the children in the cluster, even if the child is fully immunized.

Item 16 Print your name as the interviewer when the cluster form for infant immunization has been completed for the required number of children.

Note: Interviewers should leave the Total columns blank. The supervisor will complete these columns.

H.4 The cluster form for reasons for immunization failure (Form G.2)

Item 6 Record the immunization status of every child/woman that is in the cluster. This information comes from item 12 on the cluster form for infant immunization status or the cluster form for Td (or TT) immunization of women.

Item 7 If one or more immunization(s) have not been given, ask the child's mother or caretaker to give the most important reason why the child did not receive all the immunizations in the series, or ask the woman why she did not have all the Td (or TT) immunizations she should have had. This is an open-ended question. Wait until the respondent answers in her own words. Do not read the list of possible answers.

Write "X" in the box closest to the answer given. Supervisors should review and summarize the responses during the data collation stage.

Note: If it is felt that categorization or precoding of possible responses may risk missing potentially important information from the respondents, the interviewers can be instructed simply to write down verbatim the reply given by the child's mother or caretaker or by the woman (provide a suitably designed form for this method). The survey supervisors and coordinator will later review all responses and decided on appropriate categories for presentation of the analysis.

Note: Interviewers will leave the "Total" column blank. The supervisor will complete it.

H.5 The cluster form for tetanus toxoid immunization of women (Form G.3)

Upon reaching a household where there is a child aged 0–11 months, ask to speak to the mother of that child. If the mother is not available, then for the purpose of Td (or TT) coverage, go to the next household.

Item 5 If there are any resident children in the household whose ages fall within the age range for their mothers to be evaluated, record the name of the mother.

Item 6 Record the date of birth of the child. Ensure that the date falls within the range specified under item 4 in the form. If the date is not within the range, do not include the mother in the Td (or TT) portion of the survey.

Item 7 Record the total number of pregnancies the mother has had in her lifetime, including the pregnancy that lead to the birth of the child aged 0–11 months. Do not include pregnancies *after* the birth of that child. This manual will call the pregnancy that led to the birth of the child aged 0–11 months the "last pregnancy". Also include pregnancies that led to a stillbirth, or a premature birth, as well as children who may have died after delivery.

Item 8 Previous versions of the coverage survey protocol did not allow identification of doses received in the last pregnancy. This version will provide a possibility to compute these doses. The data requested under item 8 should be obtained from history. The data will be used to compute TT coverage in the *last* pregnancy, and to calculate the protection at birth. Note the following definitions of terms:

- **Number of Td (and/or TT) doses received prior to last pregnancy.** Include all Td (and/or TT) doses received *before* the last pregnancy, including doses received in supplementary immunization activities (SIA).
- **Number of Td (and/or TT) doses received during the last pregnancy**. Include all Td (and/or TT) doses received during the last pregnancy, including those received during SIA.
- **Card available for Td (and/or TT) received in the last pregnancy**. Ask whether the mother can show you a card on which the Td (and/or TT) doses received during the last pregnancy have been recorded. Mark "Y" (yes) only if the interviewer is shown the card.

- **Was a card given?** For women who have a card, the answer is yes. For those without a card, the question is to determine whether the mother lost/misplaced the card or whether she never received a card. A women who does not have a card, but claims that she was given a card upon vaccination (in the last pregnancy), should be marked as "Y" (yes). If she claims no card was given, mark "N" (no).

Item 9

Before exploring the details of Td (and/or TT) doses ever received by the mother, ask to see all available immunization cards that document the mother's Td (and/or TT) immunizations. As new immunization cards are given in each pregnancy in some countries, ask for ***all*** cards that the mother may have. At the same time, ask whether any other dose was received for which she has no card. Try to determine if the immunization was for tetanus, and if so, for which doses she has a card and for which she does not.

Determine which dose was the first dose. If the first dose is marked on one of the cards, write the date of the first dose ever received for "TT1". If, for the first dose, no card is available, write "Y" (Yes) for TT1. If no dose was ever received, write "N" (No) for TT1.

Use a similar procedure for subsequent doses. It is possible that the data will have to be taken from several immunization cards, and from history. If the mother has received more than one dose of Td (and/or TT), put "Y" (or the date if the card is available) in the "Date/Y/N" box for TT2, TT3, TT4 or TT5 (depending on how many doses she received). If possible try to verify the immunizations with records at the health centre. If the mother has not received any Td (and/or TT), put "N" in the appropriate "Date/Y/N" box for the relevant dose.

When trying to determine whether the mother has received Td (and/or TT) doses for which she has no card, it is important to also ask about doses that may have been received during immunization campaigns or SIA. These may not be on the routine Td (and/or TT) immunization or antenatal care (ANC) card but must be considered.

Also remember to ask how many doses the mother received ***prior*** to the "last pregnancy", even when a card is available for the "last pregnancy". If, for example, the mother has the immunization card with the date of the last dose only, but she says she had received three doses earlier, then mark "Y" for TT1, TT2 and TT3, and write the date (from the card) for TT4.

For each dose received ask about the source of the dose, and use the following key:

- "OUT": dose received in outreach activity;
- "HOS": dose received in hospital;
- "HC": dose received in health centre;
- "PRIV": dose received from a private practitioner;
- "NGO": dose received from a nongovernmental organization;
- "SIA": dose received in a supplemental immunization activity (campaign);
- "SCH": dose received during a school immunization campaign;
- "WCV" ("well-child visit"): dose received during a visit for child immunization;
- "OTH": dose received on any other occasion (e.g. after an injury).

If the mother does not remember where she received the immunization, mark "0" in the box.

Item 10 Ask each respondent the number of times she went to antenatal care during the last pregnancy.

Item 11 Ask each respondent the number of times she went to a health facility during the last pregnancy, except for antenatal care.

Item 12 Ask where the child (now aged between 0 and 11 months) was born, and who assisted during the delivery, and mark a "Y" in the corresponding box.

Item 13 This is to be completed by the supervisor.

Item 14 Tally all households visited, regardless of whether there was a child in the defined age range.

Item 15 Print your name as interviewer and sign the form.

Note: Interviewers should leave the "Total" columns blank. The supervisor will fill them in. See Annex I.

In the survey, you would continue visiting houses within the cluster until the required number of children in the age range to be evaluated *plus* the number of children whose mothers are to be evaluated have been located. At this point, you would have completed *one* cluster. Repeating the same process for the remaining number of clusters completes the survey.

Annex I:
Supervisor instructions for completing the cluster forms

This annex is aimed at assisting supervisors when the data are to be analysed manually but the procedures described would be just as useful even if the data are to be analysed electronically for the insight into the data analysis that would be gained when a few cluster forms are completed.

I.1 Complete the cluster form for infant immunization (Form G.1)

To complete this cluster form, the supervisor follows three steps:

- checks that the immunizations recorded on the cluster form for infant immunization are valid and revises item 13 if needed;
- determines which children were fully immunized before one year of age (item 14); and
- completes the "Total" columns.

Step 1: Check that immunizations are valid and review item 13

"Valid" means that immunizations were given when the child was the appropriate age and if the immunization is one of a series that they were given after an appropriate minimum interval of time. There are three situations that would cause an immunization to be considered not valid:

1) **If the minimum age or minimum interval between doses was not respected (note that there is no maximum age or maximum interval), immunization did not follow the immunization schedule.**

   For example, a second or third DTP or OPV or hepatitis B (HepB) immunization that is given less than 28 days after the preceding immunization should be considered invalid.

2) **If the immunization is recorded "by history".**

   All immunizations (except BCG) recorded by a "+" are not considered valid. This is because the dates of immunizations recorded by history cannot be confirmed, so it is impossible to determine whether the correct schedule was followed.

3) **For BCG, if BCG is recorded "by history" and there is no scar on the child's arm.**

   If a child's BCG immunization is recorded by history, but no scar was visible, the immunization is considered not valid.

A child's BCG immunization is considered valid in all of the following situations:

- recorded by history *and* a scar is visible;
- recorded by card *and* a scar is visible;
- recorded by card, but no scar is visible. (This is because the coverage survey is trying to determine how successful the immunization services have been at reaching people, rather than monitoring the immunological response to BCG vaccine – although existence of the scar gives some indication of good injection technique.)

Circle all immunizations that are not valid according to the definitions given above. Then check that the "Fully/Partially/Not immunized" row on the cluster form (item 13) is marked correctly. The interviewers completed this item in the field, but they were not asked to consider whether the immunizations were valid. At this point, the supervisor must revise the determination of whether a child was fully, partially or not immunized, making sure that only valid immunizations are counted.

Step 2: Determine which children were fully immunized before one year of age, and complete item 14

To determine whether a child is fully immunized, reference must be made to the national immunization schedule and to the antigens included in the schedule.

The category "fully immunized" in item 13 tells you how many children were immunized under the age of two years. By comparison, item 14 asks how many children were fully immunized before one year of age.

To identify children that were fully immunized by their first birthday, consider only those children that were recorded as "fully immunized" in item 13. You then need to remove any valid immunizations that were given after the child was one year old. Look at the birth date of the child and compare the dates of all immunizations with it.

Mark any valid immunization given to the child at more than one year of age by drawing a triangle around the date. Then place a "Y" in item 14 for each child who was fully immunized before the age of one year. Record "N" in item 14 for all other children.

Note: Immunizations given after a child's first birthday may still be acceptable and, if given properly, are protective from disease. This is particularly important with measles immunization. These immunizations are not, however, included when evaluating how well a country has met its immunization coverage target of "fully immunized children" *by one year of age*. Knowing how many, or what proportion, of the children were fully immunized (i.e. received all doses) from one year of age is important in assessing timeliness of protection.

Step 3: Complete the "Total" column

When completing the "Total" columns on the cluster form for infant immunization remember the following:

- When calculating the "Total/Card" column, only count the number of valid immunizations that were verified by the presence of an immunization card (that is, for which a date was recorded).
- When calculating the "Total/Card plus history" column, add together the immunizations that fall into the following three categories:
  a) valid immunizations verified by card;
  b) immunizations that were verified by card, but were determined to be incorrectly given; and
  c) immunizations recorded by history (that is, a "+" was recorded).

I.2 Complete the cluster form for reasons for immunization failure (children) (Form G.2)

There are two steps to completing this cluster form:

- revision of item 6 if needed; and
- completion of the "Total" column.

Step 1: Review item 6

When you checked the cluster form for infant immunization to see if the immunizations were valid, you then reviewed item 13 of that form (which stated whether a child was fully, partially or not immunized).

In this step you must review item 6 on the cluster form for reasons for immunization failure, to match any changes you made in item 13 on the infant immunization cluster form.

Note: Whenever a child's immunization status changed because one or more immunizations were determined to be not valid, you will ***not*** be able to record a reason for immunization failure on the cluster form.

Step 2: Complete the "Total" column

When completing the "Total" column of the reasons for immunization failure form (G.2), remember that every child in the cluster must be marked on this form in item 6.

If there are fewer than the number of children expected from the cluster, it is possible that only those children who were not fully immunized were recorded. If this is so, it is likely that the reason given for immunization failure does not correspond to the correct child in the cluster. Check with the interviewer of that cluster and correct the form so that all children are correctly recorded.

I.3 Complete the cluster form for tetanus toxoid immunization of women (Form G.3)

There are two steps to completing this cluster form:

- completion of item 13; and
- completion of the "Total" column.

Step 1: Complete item 13

Item 13 on the tetanus toxoid immunization of women cluster form asks whether the child was protected against neonatal tetanus. The child can only be protected at birth if the mother received the tetanus toxoid immunization at appropriate intervals prior to the child's birth. Therefore, many surveys use only data that are available by card, and do not take into account immunizations recorded by history (because the date of immunization cannot be confirmed).

When cards are available, follow steps (i) through (iii) below to find out if the child was protected against neonatal tetanus:

i) Check how many valid doses of tetanus toxoid the mother received by reviewing the minimum (there is no maximum) interval between doses and the time between the last dose given during the last pregnancy and delivery. Circle any doses that are not valid, because they were not given according to the following minimum interval schedule, or because they were recorded with a "Y". Do not take into account the circled dose.

Doses	Minimum interval
Td (orTT)1 and Td (orTT)2	four weeks
Td (orTT)2 and Td (orTT)3	six months
Td (orTT)3 and Td (orTT)4	one year
Td (orTT)4 and Td (orTT)5	one year

ii) Determine the exact period of protection after the last valid dose.

Doses	Protection
One or no dose	no protection
Two doses	3 years protection, starting 15 days after the date of the last dose
Three doses	5 years protection, starting 15 days after the date of the last dose
Four doses	10 years protection, starting 15 days after the date of the last dose
Five doses	lifelong protection

iii) Determine whether the child was protected. The child is protected if he/she was born during the period of protection identified in step iii (b) below. Put "Y" or "N" in item 13 of the cluster form.

Note: Women who produce written documentation of having received three doses of DTP in childhood should be treated as having received TT1 and TT2.

If it is decided to also include history to determine protection at birth, a different way of calculation must be adopted. In that case, all doses in item 9, whether circled or not, are considered. An example of an easy way to determine protection at birth through history is:

a) Check whether the women received two doses of Td (or TT) in her last pregnancy. If so, the child was born protected.

b) In situations where no dose or one dose of Td (or TT) was received during the last pregnancy, check if at least three doses of Td (or TT) were received any time in life, i.e. during or before the last pregnancy. If so, consider it "protected at birth". Three doses are an approximation for protection, and it is acknowledged that this may lack precision.

It must always be clearly stated whether for the calculation of "protection at birth" only data on the cards or card plus history were used.

Step 2: Complete the "Total" column

Item 8 (a,b) Calculate which dose was received in the ***last*** pregnancy.

- If no dose was received in the last pregnancy (regardless of earlier doses), tally as TT0.
- If one dose was received in the last pregnancy, add that to all previous doses, to obtain the dose received. Example: if a women had received three doses earlier in life and received one dose in the last pregnancy, tally as TT4.
- If two doses were received in the last pregnancy, tally as TT2. However, if *any* doses had been received previously, only include one dose received in the last pregnancy, as the second dose will have been invalid. Example: if a woman had received two doses earlier in life, and received two doses in the last pregnancy, tally as TT3, not TT4. The reason is that, in this case, the minimum interval criterion (i.e. 1-year interval between dose 3 and 4) has not been met, and therefore the fourth dose is not valid.

Item 8 (c,d) Add all "Yes" and all "No" answers on the questions whether a card is available and whether the card was given, and fill the relevant box. The total of yes and no answers should equal the number of women in the cluster.

Item 9 Doses need to be calculated by whether they were recorded on a card or by card plus history.

When calculating the "Total/Card" column, only count the number of valid immunizations that were verified by the presence of an immunization card (that is, for which a date was recorded).

When calculating the "Total/Card plus History" column, count:

- valid immunizations verified by card;
- immunizations recorded by history (that is, a "Y" was recorded); second doses in a pregnancy other than the first pregnancy must not be counted.

Item 9 Tally the source of the doses.

Item 10 Tally all women who had one ANC visit, and all who went twice or more. Do not include the women with two or more visits in the total of women with one visit.

Item 11 Tally all women who had one visit (other than ANC) to a health facility, and all who had at least two. Do not include the women with two or more visits in the total of women with one visit.

Item 12 Tally the births by place of delivery and by person who attended the birth.

Item 13 Tally all children protected at birth.

Annex J:

Completing the summary forms

Annex J is to help supervisors to handle the data manually.

J.1 Introductory data on all three summary forms (items 2–4) (Forms G.4, G.5, G.6)

Item 2 Write in the period of time over which the survey was conducted. Begin with the date of the first household interview and finish with the date of the final interview.

Item 3 Write the name of the geographic area in which the surveyed clusters are located.

Item 4 Write the age in months of the children who have been evaluated, or of the children whose mothers have been evaluated. (For infant immunization coverage, the age group is 12–23 months; for Td [or TT] immunization of women, 0–11 months.)

J.2 The summary form for infant immunization (Form G.4)

Item 6 Record the total number of interviews conducted in each cluster.

Item 7 Record the total number of "Yes" answers for immunization card.

Note: "No" answers are not transferred to the summary form.

Item 8 Record the number of children that have BCG immunization confirmed by card. This number is from the "Total/Card" column on the cluster form.

Record the number of children that have BCG immunization confirmed by card plus the number confirmed by history. This number is from the "Total/Card plus history" column on the cluster form.

Record the number of children that have a BCG scar that was seen by the interviewer. This number is from the Total column on the cluster form.

Count, for the whole cluster, and record the total number of times each source was indicated for BCG. Use the following figure to tally (/) each time a source is indicated:

Health centre	Hospital
Outreach	Private
NGO	SIA

After having tallied all sources for BCG, summarize the times each source was indicated.

Items 9–11 Repeat the process described in item 8 above for DTP, OPV and measles. However, when recording the total number of times a source was listed for OPV and DTP, add the sources for the three doses of each antigen together.

Item 12 Record the number of "Fully", "Partially" and "Not immunized". These numbers are found at the bottom of the "Total" column on the cluster form.

Item 13 Record the number of children in each cluster who were fully immunized before one year of age.

Item 14 Record the number of households visited.

Complete the "Total" column on the summary form by adding all numbers in each row.

J.3 The summary form for reasons for immunization failure (Form G.5)

Item 6 Transfer the information from the "Total" column on the cluster forms to the corresponding column on the summary form for immunization failure.

Item 7 Transfer the information from the "Total" column on the cluster forms to the corresponding column on the summary form. The reasons are listed in the same order on the summary form as found on the cluster form. Then calculate the subtotal for all clusters in the category "Lack of information".

Repeat the process described above for the categories "Lack of motivation" (item 8) and "Obstacles" (item 9).

Complete the "Subtotal" and "Total" columns on the summary form by adding all the relevant numbers in each row.

J.4 The summary form for tetanus toxoid immunization of women (Form G.6)

Item 5 Record the total number of mothers in the cluster.

Item 8 (a,b) Record the number of women with a specific dose of Td (or TT) in the last pregnancy.

Item 8 (c,d) Record the total number of "Yes" answers for immunization card available and received. (The "No" answers for items 8c and 8d on Form G.3 are not transferred to the summary form [G.6].)

Item 9 Record the total number of women that have TT1 immunization confirmed by card. Get this number from the column "Total/Card" on the questionnaire form (G.3). Similarly, record the total number of women that have the TT1 immunization confirmed by card plus the number by history. This number is from the "Total/Card plus history" column on the cluster form (G.3).

Repeat the steps above for TT2, TT3, TT4 and TT5.

Count the number of times that vaccine was given in the same source (OUT, HOS, ...) and enter the data in the corresponding cell.

Item 10 Record the total number of women with one ANC visit and those with two or more visits.

Item 11 Record the total number of women with one other visit and those with two or more visits.

Item 12 Record the total number for each of the delivery locations and for each of the types of delivery assistance.

Item 13 Record the total number of "Yes" answers for children aged 0–11 months who were protected against neonatal tetanus. Compute separately the ones protected based on cards, and the ones protected based on card plus history.

Note: The "No" answers on Form G.3 are not transferred to the summary form (G.6).

Item 14 Record the number of households visited.

Complete the "Total" column on the summary form by adding all numbers in each row.

Annex K:
Completing the evaluation forms

Annex K is to help supervisors to handle the data manually.

If you are analysing the data manually, or need to calculate preliminary results without confidence intervals, you can use the evaluation forms to do so.

K.1 Introductory data on all three evaluation forms (items 1 – 5) (Form G.7, G.8, G.9)

Item 1 Write the name of the geographic area in which the surveyed clusters are located.

Item 2 Write the age in months of the children who have been evaluated, or of the children whose mothers have been evaluated. (For infant immunization coverage, the age group is 12–23 months; for Td (or TT) immunization of women, 0–11 months.)

Item 3 Write the date of the first household interview.

Item 4 Write the date of the last household interview.

Item 5 From the summary form, transfer the total number of children or women surveyed.

K.2 The evaluation form on infant immunization (Form G.7)

For each of items 6–11, in the columns labelled "Number", transfer the entries from the last column of the summary form. Calculate the percentage by dividing the entry in the number column by the value in item 5 (total number of children surveyed), and multiply by 100.

For item 12, transfer the entry from the last column of the summary form, item 14. For item 13, divide the entry in item 12 by the number of clusters surveyed.

K.3 The evaluation form on reasons for immunization failure (Form G.8)

Item 6 In the "Total" column, enter the *sum* of the last column of "not immunized" and "partially immunized" from item 5 of the summary form on reasons for immunization failure. For the "Percentage" column, divide the entry in the "Total" column by the number of children or women in the survey (item 5), and multiply by 100.

Item 7 For the "Total" column, transfer the corresponding entries from the summary form on reasons for immunization failure. The reasons are in the same order as on the summary form. For the "Percentage" column, divide by the entry for item 6 (the total number of children not fully immunized), and multiply by 100.

Repeat for items 8 and 9.

K.4 The evaluation form on tetanus toxoid immunization of women (Form G.9)

For each of items 8–13, transfer from the last column (the "Total" column) of the summary form on tetanus toxoid immunization of women the corresponding number into the "Number" column on the evaluation form. Calculate the percentage by dividing by the total number of mothers surveyed (item 5), and multiplying by 100.

TT2+ in the last pregnancy (item 8) is calculated by adding all doses of TT2, TT3, TT4 and TT5 received during the last pregnancy, divided by the number of women in the survey with a child aged 0–11 months (item 5). To obtain the percentage, multiply by 100. Please note that this figure is for "Card plus history" only.

Lifetime Td (or TT) doses are calculated separately for "by card" and "by card plus history", and are usually reported separately by dose, for "by card only" and "by card plus history".

For item 15, divide the entry in item 14 by the number of clusters surveyed.

Annex L:
Analysing immunization coverage survey data

Whether the data are analysed by hand or with the help of a computer the following indicators are useful to evaluate.

Area	Indicator
Immunization card **retention**	• Percentage of children/women who ever had a card
	• Percentage of children/women with a card at the time of the survey
Immunization system **access**	• Percentage of children receiving BCG (card **or** scar **or** card and history)
	• Percentage of children receiving DTP1 (card **or** card and history)
Immunization system **utilization** (drop-out rate)	• Percentage difference in coverage between BCG and measles (card **or** card and history)
	• Percentage difference in coverage between DTP1 and measles (card **or** card and history)
	• Percentage difference in coverage between DTP1 and DTP3 (card **or** card and history)
Immunization **coverage** for maximum epidemiological impact	Percentage of: • fully immunized children • with any doses (card and history) and • with valid doses (card only)
	• Percentage of fully immunized children with valid doses by one year of age (or according to national schedule) (card only)
Integration of **new vaccines** (if monovalent new vaccines are used)	• Percentage difference between DTP1 and HepB1 (card **or** card and history)
	• Percentage difference between DTP1 and HepB1 (card **or** card and history)
	• Percentage difference between measles and yellow fever (card **or** card and history)
	• Percentage difference between DTP3 and HepB3 (card **or** card and history)
	• Percentage difference between DTP3 and Hib3 (card **or** card and history)
Invalid dose administration (adherence to schedule)	• Percentage of children receiving invalid DTP1 doses (card only) (too early)
	• Percentage of children receiving invalid DTP3 doses (card only) (too short interval)
	• Percentage of children receiving invalid measles doses (card only) (too early)
Injection technique	• Percentage of children vaccinated with BCG with BCG scar (card only)
Immunization system sex **equity**	• Percentage difference between coverage among males and females (for various antigens)
	• For Td (or TT) surveys only
Utilization	• Percentage difference between antenatal clinic (ANC) attendance and TT1 coverage
Coverage	• Percentage of women with at least two doses of Td (or TT) received during their lifetime, of which at least one dose was during the last pregnancy
Invalid dose administration	• Of those who had immunization cards, the percentage who received an invalid dose (i.e. the interval between subsequent doses was too short)

These indicators can be calculated by statistical or epidemiological software, or by hand. Results should be presented with confidence intervals which can be calculated according to the formulas below.

Lower limit[1] = $p - 1.96\sqrt{D \times \frac{p(1-p)}{n}}$

Upper limit = $p + 1.96\sqrt{D \times \frac{p(1-p)}{n}}$

where:

- p is percentage of fully immunized children
- n is number of children in the survey
- D is the design effect of the survey (if simple random sampling, use $D = 1$; for cluster surveys, generally use $D = 2$)

Example 1

A survey covered 248 children (215 of whom had cards) of whom 127 had received all the scheduled antigens within the recommended time and confirmed by their vaccination cards.

Percentage of fully immunized children with valid doses (card only)

$$= \frac{127}{215} \times 100\% = 59.1\%$$

The lower 95% confidence limit = $59.1\% - 1.96\sqrt{2 \times \frac{59.1\%(40.9\%)}{215}} = 49.9\%$

The upper 95% confidence limit = $59.1\% + 1.96\sqrt{\frac{59.1\%(40.9\%)}{215}} = 68.4\%$

[1] This is in fact an approximation of the true confidence interval; the exact calculations for a confidence interval for a cluster survey are fairly complex. Many software packages will automatically calculate such intervals for you using the correct formulas. More details on this can be found in Annex A, and in the References.

Example 2

A total of 88 children with cards were not fully immunized. Out of these, 14 gave as the reason for not bringing the child for immunization: "unaware of the need for immunization".

Percentage of children not fully immunized due to lack of mother's awareness

$$= \frac{14}{88} \times 100\% = 15.9\%$$

The lower 95% confidence limit = $15.9\% - 1.96\sqrt{2 \times \frac{15.9\%(84.1\%)}{88}} = 5.1\%$

The upper 95% confidence limit = $15.9\% + 1.96\sqrt{2 \times \frac{15.9\%(84.1\%)}{88}} = 26.7\%$

Annex M:
How to use a random number table

Choosing a random number is an important step in a coverage survey because it is the only way to ensure that there is no unconscious bias in the selection of houses and individuals to be interviewed. There are several ways to select a random number, but using a random number table is one common method. This annex gives two methods of using a table of random numbers and describes how to implement each method.

M.1 Method A

Step 1 Choose a direction (right, left, up or down) in which you will read the numbers from the table.

Step 2 Select a starting point by using one of the following methods, (a) or (b):

a) Using a currency note, select a single digit random number between 0 and 9 to identify a column. Select a two-digit random number between 01 and 25 to identify a row. (**Note**: The numbers 01–09 each count as two-digit numbers.) The five-digit number in the table that is at the intersection of the column and row you have selected is the starting point.

b) Close your eyes, and touch the random number table with a pointed object. Open your eyes. The digit closest to the point where you touched the table is the starting point.

Check that the starting point will give a number which is going to be less than or equal to the sampling interval. If not, start again before going on to step 3.

Step 3 Read the number of digits required (determined by the sampling interval) in the direction chosen in step 1. Because each individual digit in the table is random, the sequence(s) of digits can be used across spaces between the five-digit numbers. The number you end up with is your random number.

For example, let us say you decided to read numbers to the right, and by using method (b) in step 2 you identified your starting point as the number 3 in row 01, column 8 (see Table M-1 of random numbers in this annex). If the sampling interval had four digits, then your random number would be 3861. The numbers 6 and 1 come from row 01, column 9.

Note: Remember that the random number selected must be equal to or smaller than the **sampling** interval. If it is not, then another random number must be selected.

Table M-1. Random numbers table

					Column					
Row	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
01	88008	13730	06504	37113	62248	04709	17481	77450	46438	61538
02	01309	13263	70850	11487	68136	06265	36402	06164	35106	77350
03	45896	59490	98462	11032	78613	78744	13478	72648	98769	28262
04	50107	24914	99266	23640	76977	31340	43878	23128	03536	01590
05	71163	52034	03287	86680	68794	94323	95879	75529	27370	68228
06	76445	87636	23392	01883	27880	09235	55886	37532	46542	01416
07	84130	99937	86667	92780	69283	73995	00941	65606	28855	86125
08	00642	10003	08917	74937	57338	62498	08681	28890	60738	81521
09	64478	94624	82914	00608	43587	95212	92406	63366	06609	77263
10	02379	83441	90151	14081	28858	68580	66009	17687	49511	37211
										11
32525	44670	57715	38888	28199	80522	06532	48322	57247	46333	**12**
01976	16524	32784	48037	78933	50031	64123	83437	09474	73179	**13**
67952	41501	45383	78897	86627	07376	07061	40959	84155	88644	**14**
38473	83533	39754	90640	98083	39201	94259	87599	50787	75352	**15**
91079	93691	11606	49357	55363	98324	30250	20794	83946	08887	
									16	72830
10186	08121	28055	95788	03739	65182	68713	63290	57801	**17**	40947
75518	59323	64104	24926	85715	67332	49282	66781	92989	**18**	44088
70765	40826	74118	62567	75996	68126	88239	57143	06455	**19**	19154
29851	16968	66744	77786	82301	99585	23995	15725	64404	**20**	13206
90988	34929	14992	07902	23622	11858	84718	22186	35386		
								21	24102	13822
56106	13672	31473	75329	45731	47361	47713	99678	**22**	59863	62284
24742	21956	95299	24066	60121	78636	61805	39904	**23**	57389	70298
05173	48492	68455	77552	87048	16953	45811	22267	**24**	63741	76077
44579	66289	88263	54780	76661	90479	79388	15317	**25**	17417	56413
35733	27600	06266	76218	42258	35198	26953	08714			
							26	85797	58089	91501
34154	96277	83412	70244	58791	64774	75699	**27**	65145	97885	44847
37158	54385	38978	20127	40639	80977	73093	**28**	24436	65453	37073
81946	36871	97212	59592	85998	34897	97593	**29**	20891	03289	98203
05888	49306	88383	56912	12792	04498	20095	**30**	81253	41034	09730
53271	92515	08932	25983	69674	72824	04456				
						31	64337	64052	30113	05069
54535	01881	16357	72140	00903	45029	**32**	35929	76261	43784	19406
26714	96021	33162	30303	81940	91598	**33**	34525	54453	43516	48537
60593	11822	89695	80143	80351	33822	**34**	27506	45413	42176	94190
29987	90828	72361	29342	72406	44942	**35**	92413	00212	35474	22456
76958	85857	85692	75341	32682	00546					
					36	76304	57063	70591	06343	38828
15904	79837	46307	40836	69182	**37**	17680	92757	40299	98105	67139
01436	68094	78222	61283	40512	**38**	43281	36931	26091	42028	62718
38898	64356	19740	77068	78392	**39**	30647	40659	23679	04204	67628
81109	73155	68299	62768	58409	**40**	26840	42152	80242	57640	19189
47061	44640	52069	98038	49113						

M.2 Method B

Suppose you need to look for a random number between 1 and 5. Take the first number in the column labelled "1 to 5" in Table M-2, and use that. Cross it off, because it cannot be used again.

If you need to look for a random number between 1 and 8, look in the column labelled "1 to 10", and use the first random number that is not crossed off that is less than or equal to 8.

Table M-2: Random numbers by intervals

	1 to 5	1 to 5	1 to 10	1 to 10	1 to 20	1 to 30	1 to 40	1 to 50	1 to 100
1	3	1	3	7	9	4	12	36	24
2	1	1	6	8	2	19	27	4	86
3	1	3	3	1	20	28	38	5	100
4	3	3	7	7	14	24	27	17	50
5	5	1	5	8	17	15	23	33	37
6	5	3	9	1	17	22	1	1	5
7	1	3	5	1	17	10	16	50	71
8	2	3	1	4	17	6	40	13	4
9	4	3	8	1	7	27	37	41	66
10	2	3	6	3	19	1	12	22	47
11	5	2	2	10	18	15	14	12	36
12	3	4	9	3	14	24	36	8	72
13	4	2	9	10	17	24	39	11	100
14	4	2	1	2	15	2	13	39	44
15	4	5	5	9	1	2	22	27	14
16	2	3	3	8	11	5	19	15	70
17	4	2	2	4	9	20	21	37	2
18	5	2	1	8	15	14	8	47	60
19	3	5	9	5	2	14	36	7	88
20	3	5	3	7	14	6	6	43	90
21	5	5	6	7	18	22	5	41	77
22	2	2	5	1	20	22	14	1	56
23	2	2	4	6	16	19	26	8	34
24	1	5	5	6	5	3	38	35	46
25	4	4	2	1	7	9	5	11	17
26	1	1	10	2	15	8	11	37	84
27	1	1	9	1	14	5	20	20	93
28	2	4	2	8	15	18	32	17	32
29	2	1	5	6	18	30	37	9	29
30	4	5	1	6	12	9	16	4	81
31	5	3	8	4	17	28	29	21	44
32	2	4	10	1	14	30	1	32	83
33	4	1	1	8	20	12	32	15	100
34	1	1	2	10	6	2	10	6	48
35	4	4	1	9	3	15	24	36	52
36	1	4	9	9	2	11	19	36	67
37	5	1	4	1	13	21	8	27	57
38	2	2	6	2	10	23	38	8	20
39	4	3	3	6	3	6	24	11	45
40	3	3	10	10	5	26	33	40	71

M.3 Alternative: computer-generated random numbers

If you are using a computer software package (such as Excel, Stata, SPSS or SAS) there are built-in random number generator functions. For selecting your clusters in the design of the survey, these can be very helpful. For use in the fieldwork portion of the survey, you should rely mostly on printed tables such as the ones given in this annex.

References

Reference	Comment
a) Bennett S et al. A simplified general method for cluster sample surveys of health in developing countries. *World health statistics quarterly*, 1991, **44**(3): 98–106.	Describes a cluster sampling approach for health surveys that retains the simplicity of the original EPI strategy. It indicates how the effects of clustering can be taken into account when estimating means, rates and proportions.
b) Brogan D. Flagg EW. Deming M. Waldman R. Increasing the accuracy of the Expanded Programme on Immunization's cluster survey design. *Annals of Epidemiology*, 1994, **4**(4):302–311.	
c) Cochran WG. *Sampling Techniques*. Chichester, John Wiley and Sons, 1977.	Classic statistical textbook on sampling. Includes theoretical probability background.
d) Cutts FT. The use of the WHO cluster survey method for evaluating the impact of the expanded programme on immunization on target disease incidence. *Journal of Tropical Medicine & Hygiene*, Oct. 1988, **91**(5):231–239.	
e) Hansen MH, Hurwitz WN, Madow WG. *Sample survey methods and Theory*. New York, John Wiley and Sons, 1953.	Classic statistical textbook on sampling.
f) Henderson RH et al. Assessment of vaccination coverage, vaccination scar rates, and smallpox scarring in five areas of West Africa. *Bulletin of the World Health Organization,* 1973, **48**:183–194.	A paper reporting on the surveys done during the smallpox eradication exercise, which were used to design the "30x7" cluster survey methodology.
g) Henderson, RH, Sundaresan T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. *Bulletin of the World Health Organization,* 1982, **60**(2):253–260.	The original paper describing the "30 x 7" cluster survey for estimating immunization coverage.
h) Hoshaw-Woodard S. *Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage*. Geneva, World Health Organization, 2001 (WHO/V&B/01.26).	Provides a brief description of the "30 x 7" cluster survey strategy and lot quality assessment sampling. It describes the general aspect of cluster and stratified sampling designs and provides guidelines on the setting in which each of the two designs is appropriate.
i) Kalton G. Some proposed modifications for the WHO simplified cluster sampling method for estimating immunization coverage. 1998.	
j) Katz J et al. Estimation of design effects and diarrhea clustering within households and villages. *American Journal of Epidemiology,* 1993, **138**(11):994–1006.	
k) Kish L. *Survey Sampling*. London, John Wiley and Sons, 1965.	Classic statistical textbook on sampling.

Reference	Comment
l) Kish L, Hess I. On variances of ratios and their differences in multistage samples. *Journal of the American Statistical Association*, 1959, **54**:416–446.	
m) Lanata CF, Black RE. Lot quality assurance sampling techniques in health surveys in developing countries: advantages and current constraints. *World Health Statistics Quarterly,* 1991, **44**(3):133–139.	
n) Lemeshow S, Robinson D. Surveys to measure programme coverage and impact: a review of the methodology used by the Expanded Programme on Immunization. *World health statistics quarterly*, 1985, **38**(1):65–75.	Reviews and discusses the "30 x 7" cluster survey strategy for immunization coverage assessment and suggests areas needing further study to strengthen the methodology.
o) Lemeshow S et al. A computer simulation of the EPI survey strategy. *International Journal of Epidemiology*, 1985, **14**:473–481.	Computer simulation of the "30 x 7" sampling strategy, giving insights in situations under which the strategy performs satisfactorily and situations under which it is likely to fail.
p) Lemeshow S et al. *Adequacy of sample size determination.* Chichester, John Wiley & Sons (published on behalf of the World Health Organization), 1990.	The book provides solutions to typical problems and tables of minimum sample sizes for various survey and study designs. It also provides a concise exposition of the theory behind the process of determining sample sizes.
q) Lemeshow S, Stroh G. Quality assurance sampling for evaluating health parameters in developing countries. *Survey methodology,* 1989, **15**(1):71–81.	
r) Lemeshow S, Taber S. Lot quality assurance sampling: single- and double-sampling plans. *World Health Statistics Quarterly,* 1991, **44**(3):115–132.	
s) Lohr SL. *Sampling: Design and analysis*. Pacific Grove, CA, Duxbury Press, 1999.	Modern introductory-level textbook on sampling. Examples of computer code for analysing data available with text.
t) Lwanga SK, Abiprojo N. Immunization coverage survey methodological study in Indonesia. *Bulletin of the World Health Organization*, 1987, **65**:847–853.	A report of field tests of the classic "30 x 7" sampling strategy with reference to the possibility of using lot quality assurance.
u) Lwanga S, Lemeshow S. *Sample size determination in health studies. A practical manual*. Geneva, World Health Organization, 1991. [Available in French, Spanish, Chinese and Parsee.]	A practical guide to sample size determination for a variety of situations with extensive tables of sample sizes under simple random sampling.
v) Milligan P, Njie A, Bennett S. Comparison of two cluster sampling methods for health surveys in developing countries. *International Journal of Epidemiology*, 2004 Jun, **33**(3):469-476.	
w) Rea LM, Parker RA. *Designing and conducting survey research: A comprehensive guide*. San Francisco, Jossey-Bass, 1997.	Non-mathematical issues in designing surveys.
x) Robertson SE et al. The lot quality technique: a global review of applications in the assessment of health services and disease surveillance. *World Health Statistics Quarterly*, 1997, **50**(3–4):199–209.	
y) Rothenberg RB et al. Observations on the application of EPI cluster survey methods for estimating disease incidence. *Bulletin of the World Health Organization,* 1985, **63**(1):93–99.	

Reference	Comment
z) Sandiford P. Lot quality assurance sampling for monitoring immunization programmes: cost-efficient or quick and dirty? *Health Policy & Planning,* 1993, **8**(3):217–223.	
aa) Scheaffer R., Mendenhall W, Ott L. *Elementary survey sampling, 5th Ed*. Boston, Duxbury Press, 1996.	Modern introductory-level textbook on sampling.
bb) Singh J et al. Concurrent evaluation of immunization programme by Lot Quality Assurance Sampling. *Journal of Tropical Pediatric,* 1995, **41**(4):215–220.	
cc) Turner AG, Magnani RJ, Shuaib M. A not quite as quick but much cleaner alternative to the Expanded Programme on Immunization (EPI) cluster survey design. *International Journal of Epidemiology,* 1996, **25**(1)198–203.	
dd) Verma V. *The estimation and presentation of sampling errors*. London, World Fertility Survey, 1982 (Technical Bulletin No. 11).	
ee) Core information for the development of immunization policy: 2002 update. Geneva, World Health Organization, 2002 (unpublished document WHO/V&B/02.28; available on Internet at: www.who.int/vaccines-documents/DocsPDF02/www557.pdf).	
f f) Protocol for assessing neonatal tetanus mortality in the community using a combination of cluster and lot quality assurance sampling. Field test version (unpublished document WHO/V&B/02.05; available on Internet at: www.who.int/vaccines-documents/DocsPDF02/www635.pdf).	