Rapid Health Surveys

PRINCIPLES AND SAMPLING DESIGN

HANDBOOK





ICF MACRO
an ICF International Company

PUBLIC HEALTH INSTITUTE®

ACKNOWLEDGEMENTS

The authors would like to thank Ruilin Ren of ICF Macro for his expert advice and useful comments regarding sampling issues, as well as Susan Hinkins and Gary Shapiro for very useful feedback on the first draft. We are indebted to a host of previous authors who have largely contributed to the advancement of the field, and from which we have borrowed heavily, most notably Bill Weiss and Joe Valadez. We also want to thank Cindy Young-Turner for providing essential editing assistance in a very short timeframe, as well as ICF Macro's graphics and word processing team for formatting assistance and cover design. We are indebted to all these colleagues for helping make this a quality product, but any remaining mistakes are certainly ours.

Suggested citation:

Davis, R., Luna, J., Rodriguez, A., Sarriot, E. 2009. *Rapid Health Surveys: Principles and Sampling Design Handbook.* ICF Macro: Calverton, MD.

TABLE OF CONTENTS

Annexes

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

CI	Confidence Interval
DHS	Demographic and Health Surveys
EA	Enumeration Area
KIS	Key Indicators Survey
KPC	Knowledge, Practice and Coverage Survey
LQAS	Lot Quality Assurance Sampling
MICS	Multiple Indicator Cluster Survey
MIS	Malaria Indicator Survey
NGO	Nongovernmental Organization
POU	Point of Use
PPS	Probability Proportionate to Size
PSU	Primary Sampling Unit
SA	Supervision Area
SRS	Simple Random Sampling
USAID	U.S. Agency for International Development

SECTION I: INTRODUCTION AND GENERAL PRINCIPLES

INTRODUCTION

WHAT IF…

Too many public health managers, whether health district officers or development partners, seem to work "blindfolded" when it comes to understanding the health status of the populations with which they work. Development "experts" too easily blame local program managers for complacency and lack of enthusiasm. However, these "experts" fail to recognize the heavy responsibility carried by program officers and the demotivating effect that a lack of appropriate and timely information about what is really happening in the populations they serve can have on them. When timely and contextually relevant information is provided to local program officers, the response is unquestionably positive.[1] But how many of us have heard statements like the following from health managers?

> *We know how many bed nets were distributed and how many vaccines were given last year, but to tell you whether the children sleep under the net, and how many, this much I cannot tell you. I suspect that there are children who do not come for vaccinations, but I cannot tell you for sure how many they are.*
>
> *The country and the province report having made progress in promoting oral rehydration, but our situation here is different. The people here are different, and the folks in the capital have no idea what our situation is like. To tell you the truth, I can keep an eye on my staff and I encourage them to do outreach, but we really don't know what our priorities should be. We follow the national initiatives and immunization days.*

And we've all faced a moment when a project that is supposed to support district health providers starts wondering about its impact. Still too often, even with the emphasis of the last decade on "performance", such projects can provide reams of spreadsheets about each and every input of their efforts—people trained or numbers about service utilization—while knowing very little about the actual situation of the population. A statement like

> *We're waiting for the results of the last National Health Survey for our Province, but we really don't know how we compare to the Province as a whole really. It's difficult to set priorities and even more difficult to know if we've been successful,*

is still an oft-played refrain.

This manual wants to challenge this status quo.

What if we could obtain contextually meaningful information about public health indicators in a relatively short timeframe? What if we could do this without incurring high costs and in a way

---

[1] Eric Sarriot, Jim Ricca, Leo Ryan, et al. Measuring sustainability as a programming tool for health sector investments—report from a pilot sustainability assessment in five Nepalese health districts. Int J Health Plann Mgmt 23:1-25. 28 Apr. 2008. http://www3.interscience.wiley.com/journal/118903030/abstract

that would strengthen our partnership with local players? What if we could work together—districts, local authorities, and finance authorities—to examine the situation via a limited set of indicators that make sense to professionals and community members alike, that are valid and reliably measured, and then use the information gained to agree on targets for improvements? Finally, what if we could set a date, in a couple of years or so, when we could evaluate whether or not we have reached those targets? By systematically and regularly measuring meaningful indicators—using low cost methods—about population health, we would not only strengthen the evidence for our work but also build motivation and accountability with our partners.

If you share these aspirations, then this manual should come as a useful resource to you. **It is possible to obtain reliable data on population health at the local level, in a way that is timely and financially achievable for the purpose of priority setting, decision-making, performance monitoring and evaluation!** The example below provides short examples of what can be accomplished.

Example: The Information Power of Rapid Surveys
from USAID's Child Survival and Health Grants Program (data available from http://www.childsurvival.com/projects/statistics.cfm and www.statcompiler.com)

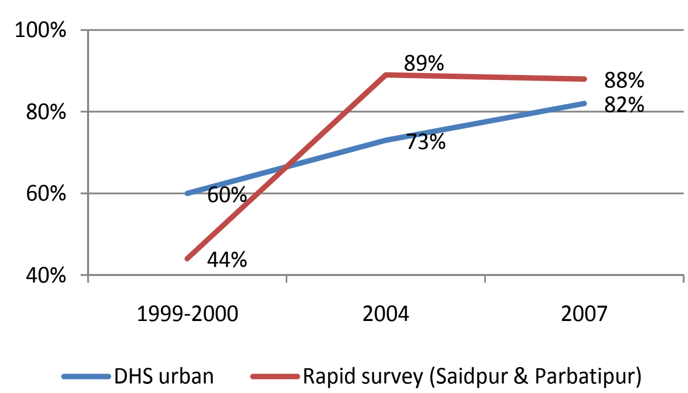
Concern Worldwide implemented an urban health project in Saidpur and Parbatipur municipalities in Bangladesh from 1999 to 2004. With the role of helping city health departments take the leadership in organizing primary health care across every neighborhood, priorities and targets had to be established. While national health coverage surveys existed, Concern needed a baseline estimate of health coverage in its operational areas. They established an estimate for key indicators, including immunization, in each of their two partner municipalities (through a method called stratifying—see below in this document) and obtained an aggregate for their entire program area. Municipality staffs were involved in the decision to carry out the survey, in the selection of modules and indicators, and in the review of findings. The estimated full immunization coverage in Saidpur and Parbatipur was 44% for 1999 [n=475], considerably below the national urban average of 60% (BDHS 1999-2000), making this a clear priority.

By 2004 the national urban full vaccination coverage had increased to 73% (BDHS). Concern and the two municipalities repeated the rapid survey and found that the estimate for the full program area was now 89% (observably and statistically different from the initial 44%). Of course, evaluation questions were raised and satisfied about the probable attribution and significance of the measured change. These data made clear that something positive had happened in the Concern intervention area.

In 2007, three years following the closure of Concern's work in Saidpur and Parbatipur, another rapid survey was conducted to assess the coverage of a series of child survival interventions, including full vaccination. While national urban estimates were now of 82% for full immunization (BDHS), the Saidpur and Parbatipur rapid survey estimate was 88% (a figure not statistically different from the previous estimate).

These data are plotted in the figure below. Even though the information presented in this box is limited, what do you think about the value of the local information provided by these rapid surveys for the municipal health departments (and, accessorily, Concern)?

Figure 1.1: Complete child immunization indicator in Saidpur and Parbatipur against national urban estimate



PURPOSE OF THIS HANDBOOK

This handbook is designed to help health managers understand the why and the how to of rapid health surveys[2]. Readers will be planners, managers or monitoring and evaluation officers working in the Mesoamerican Health Initiative and other health programs. They will be civil society organization (nongovernmental organization [NGO]) managers, leaders of government health districts, or managers of health programs implemented in regions or districts. The common need of all these professionals is for a one-stop source of information that enables them to decide whether they want to carry out a rapid survey, what they should think about in their plans, and specifically to identify and implement the appropriate sampling approach. This handbook aims to serve this need.

We have tried to adopt a direct and engaging style, being sensitive to the fact that non-specialist managers will be our audience more often than statisticians, who probably find it easier to navigate the plethora of resources already available. We have nonetheless made the effort to also provide enough technical references and expert know-how to help users deal with advanced questions in the design of a survey. However this is not a comprehensive document on sampling strategies. When we have not covered a topic in depth, we have provided extensive references and links to existing resources. In most cases, you can find these in Annex 1: Key Resources. And we have been explicit in making use of preexisting material as much as possible. This handbook focuses on the two most frequently used approaches to household surveys in developing countries: two-stage cluster sampling and then Lot Quality Assurance Sampling (LQAS). We present the most current practical recommendations on how to design and implement these approaches. These recommendations are based on decades of field experience in international health surveys. The handbook also tries to make already existing information more readily accessible to those seeking to obtain valid population-based health data at local and program levels.

Finally, while some organizations develop the internal capacity to conduct rapid surveys on their own, it is advisable to work with local survey groups and at least expert technical assistance to ensure quality at each stage of design and implementation. Even with expert assistance, however, program managers and monitoring and evaluation officers need to be actively engaged in design and implementation options and will be helped in this by this manual. Local survey groups sometimes have limited experience or no experience at all with LQAS sampling, and even less with the kind of analysis it allows. This handbook should provide enough to be of service in getting them on board. What is not advisable in any case is for program implementers to outsource and walk away from the rapid survey exercise. No matter what technical assistance is provided, organizations who commission the surveys have a key role to play in asking the meaningful questions of the survey and in bringing in stakeholders and local partners at key steps in the design of the survey.

[2] This manual deals with rapid population-based health surveys as opposed to health facility- or other institution-based survey. We will use the following terms interchangeably: rapid health survey, rapid survey, and rapid population-based survey.

STRUCTURE OF THE HANDBOOK

In Module 1, we will go through all the essential elements of planning and implementing a rapid survey. Without specifics—such as what population is surveyed and what indicators are being measured—this module will remain general in tone, but it will refer to valuable tools and resources and provide examples along the way.

The second section of the handbook is dedicated to sampling. After a brief introduction to general sampling concepts and terminology, each of the two modules of the section will address a specific sampling strategy, namely two-stage cluster sampling and then LQAS.

FIRST THINGS FIRST—WHAT ARE RAPID (POPULATION-BASED) HEALTH SURVEYS?

What we are dealing with are quantitative population-based surveys on health issues, where findings about a representative sample of a population are used to provide estimates for the entire population of study, with a known level of possible error. The health issues investigated may be knowledge about key health practices, they may be household practices, they may be availability and use of certain health services, and—in some cases—they can even be health status indicators such as the prevalence of malnutrition or anemia.[3] Collectively we will refer to these issues as "coverage," defining it generally as the percentage of people with which a program is concerned who know a key piece of information, practice a recommended behavior, or receive or use a particular service.

This definition, however, also fits national surveys, such as the Demographic and Health Surveys (DHS), funded by the U.S. Agency for International Development (USAID) and considered the gold standard in global health data, or the Multiple Indicator Cluster Survey (MICS) implemented by the United Nations Children's Fund. So what makes a survey a *rapid* survey?

The term "rapid" essentially distinguishes these surveys from national surveys—even when the latter provide disaggregation down to the level of regions or provinces—because they represent a smaller scale of data collection by:

- Being implemented at the district or sub-district level
- Being implemented within the geographic coverage area of a program
- Using simpler sampling designs
- Involving substantially smaller (orders of magnitude) sample sizes than DHS or MICS surveys
- Generally focusing on a relatively smaller number of health indicators, based on local or program priorities (we discuss this below)

"Rapid" could aptly be replaced by calling them *local* surveys, or *focused* surveys, but they are also relatively more rapid than national surveys, first because of their more narrow focus and their much smaller sample sizes, second because they often involve lower-level stakeholders

[3] In general, we recommend using rapid surveys to measure prevalence of health conditions only when direct measures—such as weight, height or hemoglobin levels—can be done accurately.

than national surveys (usually endorsed by the Minister of Health), and third because of implementation modalities that will allow a turnaround time—from design to production of information—of a few months to a year. This last characteristic is one of the most important ones, since it allows timely availability of information for decision-makers.

This also means that rapid surveys represent a different level of effort in terms of human and financial cost. Because of the simpler sampling design and shorter questionnaires, we can train interviewers and supervisors more quickly. Data analysis is also simpler and can be completed much more quickly with limited expertise. All of the above characteristics make rapid surveys appropriate as a tool to provide local program decision makers with vital information upon which they can make decisions and judgments about their efforts. Table 1.1 offers a summary of similarities and differences between large national surveys and rapid surveys.

Table 1.1: Comparing national and rapid health surveys

	Rapid survey	National survey
Stakeholders	Program and local level (district and sub-district) authorities and partners	National and regional authorities and partners
Sample size	Hundred(s)	Thousands
Power	Modest: Sub-sample analyses and disaggregation of results (i.e. sub-district) require specific adjustments (see sample size, parallel sampling).	High: Allows sub-sample analyses and disaggregation at regional/provincial levels. Can show relatively small changes in national and regional indicators.
Knowledge practice coverage indicators	Yes	Yes
Mortality data (and rare events)	No	Yes
Provides regional and national estimates	No	Yes
Provides district, sub-district, and/or project level estimates	Yes	No
Time from decision to report	5–10 months	12–36 months
Range of cost	$8,000–$25,000[4]	> $150,000[5]
Type of surveys	KPC KAP	DHS MICS

We are now ready to discuss the planning and implementation of a rapid survey through Module 1, which starts with the most basic question: why do we want to conduct a survey in the first place? Please note, many of the sections in Module 1 have been adapted from the *Knowledge, Practices and Coverage Survey 2000+ Field Guide*—see Annex 1: Key Resources for link to this document.

[4] See Section *How Much Will the Survey Cost?* This range is based on common practice in low income countries; it depends mostly on sample size, number of survey domains and logistics (e.g., travel time, distances, etc). Collection of biomarkers and other adjustments would obviously drive the cost up.

[5] The upper limit is difficult to set, considering the country and population size, the level of disaggregation required, the number of indicators, the measurement of rare events, the collection of bio-markers, which can bring the cost of a national survey to over the half a million dollar level.

MODULE 1: ESSENTIAL PRINCIPLES FOR PLANNING AND IMPLEMENTING A RAPID SURVEY

WHAT ARE RAPID SURVEYS ALL ABOUT?

WHY DECIDE TO CARRY OUT A RAPID SURVEY?

Let's start with an example[6]—

> *A health program in an African country was helping to decrease diarrhea morbidity through encouraging point of use (POU) water treatment. A target for improved POU water treatment activity was to achieve 26% of the population in the program area using POU by the end of the first year. The program conducted a rapid survey and found that for the entire program area the target of 26% had been reached. This was good news. Then looking closer, the program team found that although four out of the five sub-divisions within the program area managed by different supervisors (called supervision areas)*[7] *had reached the target, one had not. Program staff became concerned and wanted to know what was different about this supervision area from the others. They visited that supervision area, talked to community members and found out that some improvements had been made to community water systems (although not enough to reduce diarrhea) and because of these improvements community members did not think POU was necessary. With this information the project changed approach for that supervision area, which helped those communities better understand the benefits of POU.*

The foregoing example describes one—real world—example of a simple use of rapid, population-based, health surveys. It illustrates one common use of such surveys—in this case to assess progress towards a program's objectives. Other applications are discussed below.

The main reason for implementing rapid health surveys is one of accountability, particularly as local information about population health is likely to be unavailable to health practitioners, whether they work for the government (health district) or a non-governmental organization program or project.

In the current era of performance-based funding, we certainly have to hope that performance will more and more be defined by population-based outcomes rather than by financial inputs engaged and basic service or program outputs. We would be delighted if this manual encouraged this shift.

The old adage is that "if you can't measure it, you can't manage it." Yet, many district officers are being asked to do just that, having only service statistics and national-to-regional level indicators to inform their activities. And project managers are very much—for the length of their project—in the same situation: they have to make management decisions and need the pertinent

[6] Adapted from Medical Teams International. Presentation by Todd Nitkin at USAID Bureau of Global Health Mini University 2008.

[7] Supervision areas will be discussed later in the handbook in sections on LQAS.

information to do so. As we see in Table 1.1 of the Introduction, national surveys usually do not provide information below the regional (or province) level and take some time (in some cases years) before results are available. Both the project manager and the district officer mobilized to improve the health indicators at the local level will need to have valid indicator estimates to decide what the priorities are and how much work is going to be required. The lack of local information on their population is probably a major demotivating factor for many health planners. And service statistics, as important and useful as they are, simply do not provide equivalent information.

Because rapid surveys are based on a random sampling approach of all the local population, they provide a true population estimate. Service statistics are sometimes used to suggest coverage figures, but they have to rely on ratio of services delivered per estimated total population and can be subject to aberrant results (such as coverage rates more than 100 percent). This is not the place to discuss the important value of routine information in monitoring activities and services, but *rapid surveys—when implemented properly—provide health managers with something far different: a true population estimate of selected health coverage indicators.*

THE IMPORTANCE OF RAPID SURVEYS IS WIDELY UNDERAPPRECIATED

The importance of this local information on population health indicators is still greatly underestimated. Important projects, regardless of sponsor (multilateral, bilateral, private foundation, or global partnership) still too often fail to provide the most basic (locally relevant) evidence for setting up objectives or the demonstration of progress made toward targets.

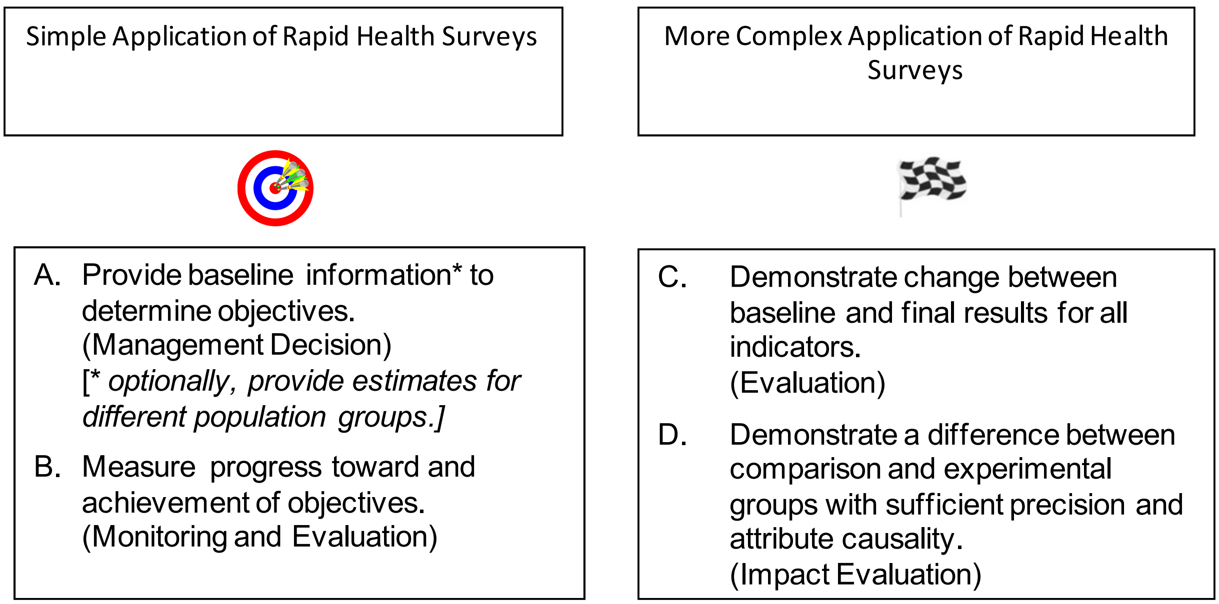
The value of rapid surveys gets lost in discussions about their relative weaknesses compared with large sample (national) surveys. As Table 1.1 shows, the survey has to be tailored to its purpose. National, large sample population surveys have specific purposes, and so do rapid surveys.

It is very important at this stage in global health efforts to discuss the value of rapid surveys based on the specific management information they provide and the conditions for their valid use. The next section will discuss what rapid surveys can be used for, and the rest of the document will present how they should be planned and implemented to meet their aim. Based on these elements, health practitioners should, much more frequently than they do now, take the time to determine the value of essential health coverage indicators at the local level where their interventions will be implemented, then plan and implement surveys that can provide critical local-level information, recognizing that such surveys are simply different in nature, purpose, and cost from large-scale surveys, even if they share indicators and essential principles for validity. It is also important to not see rapid surveys in isolation, but as one of several complementary sources of health information needed for decision making and monitoring and evaluation of projects.

DEFINING USES FOR A RAPID SURVEY

The POU example above could be considered a monitoring (process evaluation) activity since it comes at the end of year one. However, we could use a rapid survey to learn about POU at the end of the program when we are formally assessing its overall value (summative evaluation) and obviously, we could use it at baseline to learn about POU practices before we even begin activities—perhaps to determine whether such activities are even necessary (formative evaluation). Let us consider one by one the main objectives of health practitioners requiring them to implement a rapid health survey. Figure M1.1 below suggests that some of these objectives will be easily met by the standard application of basic rapid survey design, while others are often beyond the information capacity of rapid surveys, at least not without some important adjustments.

Figure M1.1: Use of rapid surveys in practice[8]



A. A health manager wants to <u>establish baseline estimates</u> for key indicators and define the most critical needs in a particular area in order to implement a program.

This is what is called "formative evaluation" and is an essential step in establishing a program or a plan based on evidence about the local situation. We will see that the approaches presented in this handbook will readily provide the information sought in this case.

A survey implemented at this initial phase can obviously provide the baseline for project interventions. Rapid population-based surveys are mostly used to measure coverage indicators such as knowledge (for example caretaker knowledge about proper child nutrition behaviors), practices (such as use of condoms), or service access (for example measles vaccine coverage). In some cases, a survey could be designed to provide information about the prevalence of certain health conditions or to quantify the existence of specific risk factors for a pathology in a given community. However, especially for low prevalence conditions, this usually requires a large sample size in order to obtain a valid disease prevalence estimate, which increases the time, effort, and budget for implementing the survey. As for risk factors, the survey will only serve to quantify the prevalence of these factors in a community. Associating these factors with disease prevalence will usually require the use of other investigation approaches.

[8] Adapted from Sarriot E., Winch P., Weiss W., et al. Methodology and Sampling Issues for KPC Surveys. July 2009. Johns Hopkins University, School of Public Health, Department of International Health. http://www.childsurvival.com/kpc2000/kpc2004.cfm

Moving toward a little more complexity, a district officer or program manager may want to compare the level of key indicators in different population groups (i.e. rural vs. urban residents). This is commonly done through rapid surveys and this manual will explain how this can be done.

B. A health project or health district wants to assess the progress they are making toward set objectives for access to health services and for improving health outcomes.

This is the domain of monitoring, also called "process evaluation," where a project is going to establish how much progress has been made toward a defined objective. The main thing here is that we are trying to establish an estimate for key health coverage indicators and whether, yes or no, they have reached a pre-set target level. Here again most rapid surveys based on appropriate sampling methods and sizes will readily provide this information with sufficient statistical "power."[9]

C. To assess the overall value of what they have been doing for a period of time.

This use of a survey is part of a "summative evaluation" (mid-term or final). At a minimum, there is value in establishing whether project's targets have been achieved. Too many practitioners are discouraged from documenting such a basic result because:

- They fear that a health survey is beyond their capacity and resources (we hope this manual will help alleviate that fear).
- Researchers have convinced them that there is no value in demonstrating that targets have been reached unless there is a change from baseline (statistically significant) and causality can be proved with full scientific rigor.

  While this is undeniably the gold standard and is essential, for example, to the introduction of new therapies (e.g., zinc as a treatment of diarrhea), managers are not in the business of efficacy trials. When a set of proven interventions is effectively implemented in a community, and a sound evaluation methodology accounts for events which may have contributed positively or negatively to certain indicators, a manager will make reasonable decisions based on whether indicators have reached the desired targets or not.

Still, a project may want to show that the final value of an indicator after implementation is statistically significantly higher than it was at baseline. Showing such a difference will call for more statistical power and consequently larger sample sizes for surveys. Rapid surveys will only allow this if proper care is taken to ensure it. We discuss this issue in Section II.

---

[9] The power of a statistical test is the probability that it will reject the hypothesis that an estimate for an indicator is not different from a given target value (null hypothesis) when it fact the indicator is above or below the target. The larger the sample, the greater the power or ability to claim as statistically significant an observed difference between an estimate and a target number.

D. Finally, an ambitious health planner may want to actually demonstrate the impact of an intervention.

Without going into details about evaluation design and strategies, a typical way of demonstrating impact (improvements caused by an intervention) is to use comparison if not control areas (non-intervention) and demonstrate that a pre-post difference observed in a district of intervention, for example, is statistically significantly greater than the difference observed in a non-intervention but otherwise comparable district. In principle, nothing prevents the use of rapid surveys for this purpose, but most will not achieve the level of complexity and the power required to carry out impact evaluation.

In the rest of this module, we will mostly address the first three uses of rapid surveys—for baselines and for comparing results at any stage of the project to stated targets—but provide relevant information in case a local project intends to, for example, demonstrate changes between a baseline and a mid-term or end-of-project status.

FURTHER CONSIDERATIONS ON THE PROPER USE OF RAPID SURVEYS

It is important for health planners to realize that rapid surveys are only one part of the overall process of identifying needs, implementing programs, tracking progress, and assessing value of achievements. Planners need to consider what other sources of data can be tapped into before committing to a population-based survey. Bearing this in mind, it is worth reminding the reader of five important principles that we all need to consider in relation to the use made of rapid surveys.

Principle 1: Plan first, then measure.

Rapid surveys can be used to help plan a health program by helping identify key priority interventions or actions. Even then, planners need to first ask which indicators they want to examine and compare. There is no pre-made survey off the shelf able to answer these questions.

Additionally, without a clear plan it is difficult to use survey data for monitoring and evaluation. Setting clear objectives and naming the indicators one will use to assess progress toward their accomplishment is, therefore, an essential preliminary step to planning one or more surveys.

Principle 2: Organizational (or program) learning agendas are about more than survey data.

Rapid surveys, while requiring focused attention and skill building, are not the only or even the most important part of an organization's learning agenda. They can provide input into decision making, but they cannot make decisions. They can provide useful data, but they cannot (alone) provide answers about directions to take. They are part of an overall commitment to learning with and by all stakeholders. Additional methods, such as qualitative data collection (e.g., in-depth interviews with participants, focus groups, observation), participatory learning activities, and budget analysis, to name a few, are also part of an organization's learning agenda, providing critical inputs that must be processed by stakeholders and used—however imperfect and incomplete the data—to make decisions.

One of the great merits of rapid surveys is that, precisely because they are local and relatively rapidly implemented, they can actively involve local stakeholders from the design stages (at least in defining what indicators will be examined) to the review of findings and ensuing programmatic decision-making. Providing local stakeholders with local and recent data to engage in setting priorities and establishing plans has proved to be a highly empowering participatory step. When used for monitoring and evaluation, the repetition of surveys with local stakeholders and the review of findings through participatory forums have shown to have substantial value, from building ownership and accountability to mobilizing partners from different sectors of society.

Principle 3: Numbers are not magical.

Rapid surveys are designed to yield numeric information and, for that reason alone, are often viewed as more valid, more real, or having more explanatory power than other assessment methods. However, it is relatively easy to generate numbers. It is harder to ensure that the numbers actually represent a meaningful and honest assessment of what is really going on in a population. The standards set forth in this manual are given to maximize the probability of making valid inferences about what is going on in a population, but health planners need to be aware that there are many places in which errors can be introduced. Numbers alone thus provide no guarantee that an accurate description is being given.

Principle 4: Rigor is critical in all approaches ...

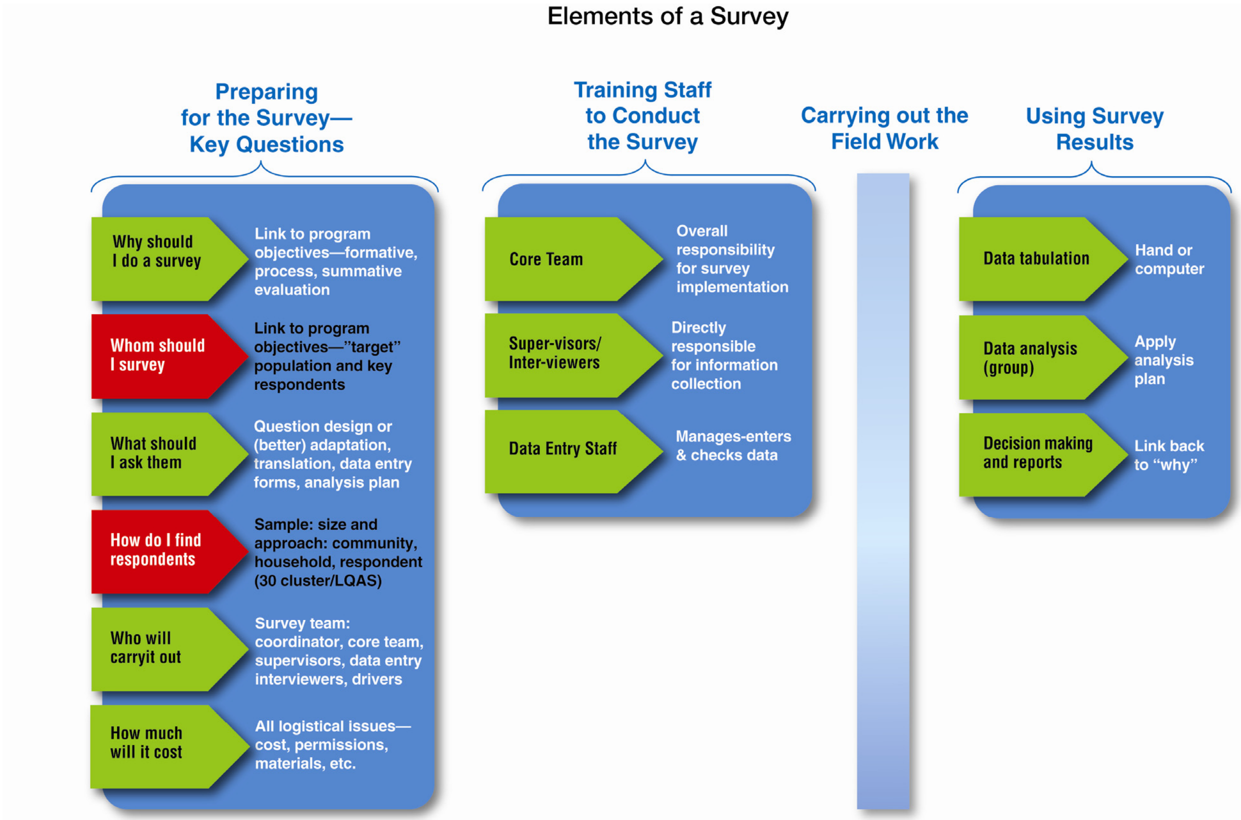
This handbook defines standards for the rigorous conduct of rapid surveys. However *all* methods require a level of rigor analogous to these. This handbook does not describe the necessary rigor needed for other approaches, notably qualitative approaches, but it does provide some references on the subject. However, as we have seen that numbers do not tell the entire story (Principle 3), program implementers will need to ensure that complementary studies—quantitative or qualitative—are implemented with equal rigor.

Principle 5: ... but quantitative methods do require special attention.

Through quantitative surveys we are using the observations of the few in order to make statements about the whole. In other words, we use a relatively small sample of a population (children under five or women of reproductive age, for example) in order to make statements about the entire population of interest. Surveys that generalize from the few to the many require that close attention be paid to issues such as standardized question design and administration, and randomness in selection of respondents throughout the entire process. For this reason, quantitative methods *do* require special attention so that we can be confident in the conclusions we draw. (But don't neglect Principle 4 because those more qualitative approaches seem more approachable.)

With this background and these principles in mind, we turn now to a map of the rapid survey process that we will follow in this manual.

Figure M1.2: Road map to a valid rapid survey



PLANNING A RAPID HEALTH SURVEY

The foregoing lays out a brief summary of what a rapid (population-based) survey is for. Of course surveying is not as simple as merely going out and talking to people if one is concerned with using a subset of the population to draw conclusions about the whole. Careful preparation is necessary to ensure that the approach to asking questions is standardized (the same for everyone), that rigorous rules are applied to sampling (selecting the subset), and that everyone involved has a clear sense of his or her role. This is why, for rapid surveys, some of the most important and time-consuming work is done well before the survey (see the Tale of the Humpty Dumpty Survey box). After the initial preparation, one must carefully train all field workers who will actually ask the questions and those who oversee their work. Of course, conducting the survey itself is labor intensive, but this work is then followed by additional, and critical, work to tabulate, analyze and draw conclusions of the data generated by asking questions. Thus, the actual field-based work of a survey is just one part of a broader set of activities.

Figure M1.2 lays this out in terms of four broad areas, each of which requires careful planning: 1) preparing for the survey; 2) training staff to conduct the survey; 3) carrying out the field work of the survey and; 4) using the survey results.

The remainder of this module walks through the elements of this diagram to lay out an overall understanding of what is required to conduct a rapid survey. However, it is important to note that each of the elements in this diagram represents broad areas for which much detail, which cannot be provided in this brief manual, exists. Therefore, this introduction will summarize briefly the major issues related to each one. You will find resources that you can use to go much deeper into each topic in Annex 1: Key Resources. The exceptions to this are the two boxes and text in red text in Figure M1.2. These boxes concern the issue of sampling—interviewing the few in order to draw conclusions about the whole mentioned above. The issue of sampling is one of the most challenging in conducting rapid surveys and thus the majority of this manual deals with sampling in great detail. Therefore, we will mention sampling in passing in this section and treat it in detail in Section II.

Tale of the Humpty Dumpty Survey

When it comes to implementing a rapid survey, some steps come with the reputation of being complex while others are considered easier. Unfortunately, those simpler steps can have as great an impact on the overall value of the survey as those feared by the average practitioner.

Sampling, data processing, and analysis are three areas where programs will rapidly call on outside help, while deciding on the types of indicators and questions, drafting a questionnaire, and asking questions of women are often considered much simpler tasks. During implementation, however, non-sampling errors (coverage errors, errors committed in survey implementation and data processing, non-response, etc.) are usually the most important sources of error.

It is a common risk and mistake in the implementation of rapid surveys to pay insufficient attention to the "simple" steps. Types of mistakes are: planning for a survey too late to establish a valid baseline, failing to quite simply check that the indicators selected actually correspond to the objectives of the project, developing insufficiently tested questions which lead to variation in the response, insufficient training and supervision of surveyors who can be prone to introduce bias in the responses, "making up" indicators which will not allow comparisons, and confusing the colloquial meaning of "random" for its stricter statistical sense, thus violating the random sampling principle essential to valid survey results.

These mistakes can lead programs to call for evaluation and survey experts, biostatisticians, and epidemiologists to come work with a dataset and extract meaning from it. But if the essential principles laid out in this chapter are violated, ***all the king's men and all the king's horses, even the king's statisticians and consultants won't put back together a survey which violated fundamental principles in design.***

A key message of this manual is: "rapid doesn't mean sloppy". Rapid surveys can be immensely valuable, and they are feasible at a reasonable cost. However, each step in the design and implementation of a survey, even the simple ones, needs to be followed with rigor from the start. Valid local information is priceless, but misleading or meaningless information is too expensive at any price.

PREPARING FOR THE SURVEY—KEY QUESTIONS TO ASK

Why should I do a survey?

The first step of preparing for a survey is, not surprisingly, to determine why it is important to do one at all. We have introduced this topic above but want to emphasize a few important points here. This handbook is designed for field practitioners who are implementing some kind of program in a defined population to improve the health of the population. Thus, the rapid survey you are conducting is, presumably, related to the objectives you want to accomplish in your program. You may need to define the focus of the program and set objectives. Or you may need to assess the progress you are making toward objectives as you go along. Or, you may need to make a judgment about the program at some critical juncture to determine whether/how it should continue or be replicated or expanded. It is important to distinguish these program-based approaches from research approaches that may use similar tools (questions, sampling, etc.) in order to test hypotheses or demonstrate causality about certain interventions.

The following figure places population-based (rapid) surveys in the context of broader program planning and implementation. You will notice from this figure that rapid surveys are just one part of an overall program planning and implementation process and only one among several tools that an organization might use to "learn" in order to set objectives, assess progress or make judgments.

Figure M1.3: Key program/project planning, learning, and management tasks

Key Program/Project Planning, Learning, and Management Tasks

Phase	Preparatory "Diagnoses" and Assessments	Detailed Implementation Planning	Implementation	Evaluation
Purpose (Including Learning)	To prepare for program development and action by better understanding— ▪ The social/economic/ environmental setting ▪ The needs and assets of various potential stakeholders ▪ Specific sectoral (health, education, etc.) challenges ▪ Key practices and behaviors to be addressed.	To develop a clear plan describing in what specific ways the organization(s) will act to deal with the problems/challenges identified during the preparatory phase—detailing key resources, expected outputs and outcomes, and key assumptions.	To implement the plans and assess, continuously, the fidelity of implementation (per the plan in terms of quality and timeliness) AND identify potential changes. Learning here includes stakeholder and participant views.	To assess the immediate effect(s) of the actions (program or project) on desired or planned outcomes—including unintended outcomes.
Learning and Planning Approaches Used	▪ Key informant interviews ▪ Group-based/Participatory tools ▪ Population-based surveys ▪ Observation (including participant) ▪ Secondary sources (documents, reports, literature) ▪ Institutional assessments	▪ Problem tree ▪ Objective tree ▪ LogFrame matrix ▪ SWOT analysis ▪ Gantt chart ▪ Budget ▪ Information system	▪ Structured observation ▪ Supervision visits/discussions ▪ Participatory tools ▪ Mini-surveys (population-based) ▪ Key informant interviews ▪ Record reviews ▪ Audits/Bookkeeping	▪ Population-based surveys ▪ Client/Member satisfaction surveys ▪ Group-based/Participatory tools ▪ Key informant interviews ▪ Institutional assessments
Resources Needed to Apply Approaches	▪ Semi-structured/Standardized interviewing techniques ▪ Group facilitation/Interviewing skills ▪ Sample survey methodologies (sampling, survey design, structured interviewing, analysis) ▪ Observation tools	While no specific "research" tools are listed here, some form of research (literature reviews or field-based methods) may be necessary to support the problem and objective trees.	▪ Semi-structured/Standardized interviewing techniques ▪ Group facilitation/Interviewing skills ▪ Sample survey methodologies ▪ Observation tools ▪ Feedback techniques	▪ Semi-structured/Standardized interviewing techniques ▪ Group facilitation/ Interviewing skills ▪ Sample survey methodologies ▪ Observation tools
Additional Key Terms	Some refer to this as— ▪ Formative Evaluation ▪ Analysis Phase ▪ Project Design Phase. See also Action Research, Appreciative Inquiry, and their links to broader strategic planning.	Planning process at this phase is referred to as logic model, results framework, logical framework etc. The terms used to develop the LogFrame (in particular) vary.	Some refer to this as— ▪ Progress Tracking ▪ Monitoring ▪ Implementation Evaluation ▪ Process Evaluation. See also Quality Improvement. Involves inputs and outputs.	Some refer to this as— ▪ Summative Evaluation ▪ Impact Evaluation ▪ Outcome Evaluation. See also Most Significant Change (MSC). Involves outcomes and impacts.

Notice a few things about this figure: First we can use population-based (rapid) surveys before a program begins to assess needs and develop the program. In this usage, rapid surveys provide information needed to specify needs more precisely and gain an understanding of the broader context in which you will carry out your program. Such information can and should feed into detailed implementation planning. In this phase, a program uses rapid survey results and results from other diagnostic tools to develop a clear plan with clear objectives. This is important because a program's further use of rapid surveys (going back to the question with which we started this section) is for the purpose of assessing progress toward these objectives—be it during implementation or in some final way.

As noted above, we commonly refer to these three uses as formative evaluation (or needs assessment), process evaluation (or monitoring), and summative evaluation. Notice that the resources (skills, actions) needed to conduct rapid surveys for these three purposes are the same. Thus, the skills developed to conduct rapid surveys for needs assessments are the same ones needed to conduct surveys for monitoring and evaluation purposes. Obviously the specifics (questions asked for example) might change, but the overall method, including the use of structured interviewing, drawing a sample, finding respondents, conducting interviews, and tabulating and analyzing results, is the same. Notice finally that the rapid survey is only one among several tools to assess need, monitor progress, and evaluate the value of an effort. This manual does not deal with these other approaches, but they are critical to an overall organizational learning agenda as we noted in the key principles above.[10]

Whom should I survey?

This question begins to get at the issue of sampling—using information about a subset of a population in order to make statements about the whole. However, here we are dealing with the question of who in relation to program objectives. It is especially important in health programs to consider who the focus of the program is and who is able to give you information you can use to assess the program. The two may not be the same, and the question of whom to survey becomes important. For example, your program may focus on the nutritional status of children, and your program objectives may include something about improving their status. The specific objectives of your program might deal with diarrhea, quality of food, or breastfeeding, for example. Obviously your population of interest is children (perhaps all children under five years of age—the most vulnerable to becoming malnourished). However, it is not reasonable to directly survey children (though you may measure their weight or height). In such cases you are going to want to interview caretakers of these children--the mothers or fathers or others who are most likely to engage in behaviors that affect the health of the children. Thus, while the objectives concern children, you may need to survey others whose behavior is critical to the objectives you are seeking.[11]

Further, in certain programs, objectives may concern a variety of groups, and the question of whom to survey must account for multiple groups. We shall see how this makes sampling and surveying more challenging in the sections below. For example, an HIV/AIDS prevention program might focus on various groups in a population, especially at-risk groups. Thus a given program may focus on injecting drug users, commercial sex workers, teenagers, or women of reproductive age (due to the problem of mother-to-child transmission). While the program

[10] Much has been written about the process of developing a monitoring and evaluation system (including needs assessments). Specifically, there is a large body of literature on the process of developing program goals, objectives, outputs and inputs and linking them in a logical framework. One useful document is the PVO Child Survival and Health Grants Program: Technical Reference Materials on Monitoring and Evaluation. Another is Medical Team International's Monitoring and Evaluation Manual. We encourage you to consult these documents to examine the steps for implementing a full monitoring and evaluation system. (The links to both of them and to others on the issue of project planning are in Annex 1: Key Resources.)

[11] Throughout this manual we will refer to a survey developed by a group of organizations working in child health with support from USAID called the Knowledge, Practice and Coverage (KPC) survey. This survey is designed for use in health programs that focus on both maternal and child health issues but the respondents for nearly all the questions are the mother of children 0-23 months of age. We will discuss this particular survey more below

concerns preventing the transmission of HIV/AIDS, it may be necessary to ask different groups different questions related to how the program addresses their particular needs.

What should I ask them?

Again, this question is very much related to program objectives and concerns what I *need* to know—and how many questions I must ask. A key concept here is operationalization. This word merely means that in asking questions we need to start with *program objectives* develop some *indicators* that will enable us to assess, in simple but acceptable terms, whether we are achieving the objectives and then develop one or more *questions* or *observations* (from the population) to fully measure the indicators. This entire process is operationalizing concepts into specific questions that we can ask. Moving from objectives to indicators is described as follows:[12]

> *An indicator... measures one aspect of a program or project that is directly related to the program's results or objectives. The value of an indicator changes from baseline to the time of the evaluation. An indicator presents this change in a meaningful way such as a percentage or number. Indicators are like clues, signs or markers that inform us on whether or not the program is achieving its results or objectives...*

Indicators need to be:

- Valid (an accurate measure of a behavior, practice, or task)
- Reliable (consistently measurable, in a the same way, by different observers)
- Measurable (quantifiable using available tools and methods)
- Precise (defined so people are clear about what they are measuring)
- Programmatically important (linked to achieving results or objectives need to achieve impact)
- Comparable (can be compared across different target groups or program approaches)

Once we have defined indicators that have these characteristics, we can proceed to develop questions that help us compute the indicators. An example of this operationalization process is useful. A goal of typical child health program may be to improve the health and nutritional status of children under the age of two. This goal might lead to an objective (or result), such as improve the health seeking or care giving practices of care givers of children under the age of two. Needs assessments may reveal that an important way to do this in a program is to help caregivers more effectively prevent diarrhea and treat it when it does occur (e.g., via rehydration). To this point we have stated a series of objectives and sub-objectives or results. One way to operationalize the sub-objectives related to diarrhea is provided in Table M1.1.

[12] *PVO Child Survival and Health Grants Program: Technical Reference Materials on Monitoring and Evaluation*—see Annex 1: Key Resources link to this document.

Table M1.1: Example of operationalization of an Indicator

Indicator	Percentage of children aged 0-23 months with diarrhea in the last two weeks who were offered more fluids during the illness.
↓	↓
Questions* you need to ask in order to compute this indicator:	1. Has (name of child) had diarrhea in the last 2 weeks? If the mother says "no," then it will not be possible to ask her questions about what she did. However, if she answer "yes," then you would ask the following two questions: 2. When (name of child) had diarrhea, did you breastfeed him/her less than usual, about the same amount, or more than usual? 3. When (name of child) had diarrhea, was he/she offered less than usual to drink, about the same amount, or more than usual to drink?

* As noted above, you would be asking this question of a caregiver (in this case, the mother) who has a child between 0 and 23 months of age.

So, to compute the indicator, you would use the answers to the questions in the following way, which would give you a proportion (and, which multiplied by 100, gives you a percentage):

- Numerator: Number for which the answer was "more than usual" to either Question 2 or Question 3
- Denominator: Number of children whose caregivers said they had had diarrhea in the last two weeks

Notice a few things about this example:

- The overall program objective and the sub-objectives are not merely about diarrhea. Thus, the indicator that concerns what happened during a recent episode of diarrhea is only one among (perhaps) several indicators.
- The objectives concern children, but the questions and the behavior indicate that a child's caregiver is the appropriate person to ask.
- It takes three questions to calculate a single indicator.
- In this case, if the answer to the first question is "no," then it is not possible to assess the behavior of that particular caregiver in relation to the indicator. This will be important when we later consider the question of how many people you must survey in order to draw conclusions.

This may seem like a long and complex process. There were many steps between naming a desired outcome to developing a question which, only partly, helps you assess whether you have achieved the objective. The good news, at least in health programming, is that much of the work of moving from objectives to indicators to questions has already been done. Thus, the process of deciding what to ask may be less about developing new indicators and questions and more about adapting what is already out there. Surveys such as the KPC, or the Measure Evaluation Guide for Monitoring and Evaluation of Child Health Projects provide a useful set of standardized indicators and questions that fully operationalize them in ways that enable a program to measure them with validity, precision, and reliability. We strongly suggest that, when possible, programs draw on standardized indicators and questions. An additional benefit is that standardized indicators will be more readily comparable with national statistics, such as those produced through the DHS. You can find links to these and other standardized surveys in Annex 1: Key Resources.[13]

Despite these resources, there may be instances when you find you must develop new questions. The design of questionnaires is a specialized and time-consuming proposition, and we recommend you consult Annex 1: Key Resources. Several resources referred to in the annex deal with question design. Use them as you consider the need to develop new questions.

Finally, the question of "what to ask" also points to another important issue in rapid surveys. For any data collection activity, it is very important to distinguish between information that is "needed" versus information that is "interesting" or would be "nice to know". For rapid surveys, especially, it is critical to concentrate on a limited number of indicators critical for decision-making or for action. Typically rapid surveys should take less than 45 minutes to conduct with each respondent.

How do I find respondents?

We have already begun to answer this above as we discussed the issue of whom to survey. However, even if you know the kind of person you want to interview—for example, caregivers of children 0–24 months—you still have a number of important issues to deal with. All of the following questions are critical to the issue of finding respondents. We will ask them in general terms here and answer each one in detail in Section II.

a. ***How many respondents do I need to interview?*** This question gets back to the challenge of using a subset of a population to say something about the entire population. So, in the example we have been using, if our program (among other things) is trying to improve the way caregivers deal with diarrhea, then I want to know how well my program has done to change behavior of all the caregivers of children 0–24 months of age in the area where I am carrying out the program. As noted previously, it is not possible to interview every caregiver, so I use a sample—a subset—of all such caregivers in order to draw conclusions about the whole. So, the first question about finding respondents is, "how many do I need to find?

b. ***Where do I go to interview the respondents?*** There are actually a number of questions here, but they all get to the same point—how do I actually decide whom to interview? We know that the respondents of interest live in all the communities in our program areas—towns and villages big and small scattered over a (potentially) large area. So, the simple answer to this question is, I go to the communities where respondents live and I interview them there. This would be correct, but there are still challenges. Unless the program has found a way to register every caregiver (and keep the list up to date over the life of the program), it is still not clear how to go about finding them. A simple way

13 The KPC survey, in particular, provides detailed information about standardized indicators and the specific questions needed to compute them. Further, it describes, in detail, the exact numerators and denominators—based on specific questions—needed to compute the indicators. It also provides examples of how to format questions to facilitate the work of those asking questions and those entering data for analysis. Thus, many of the issues concerning "what should I ask them?" have already been dealt with in detail.

would be go to a nearby community and find all the caregivers and ask them. We could do this until we reach the total number we have decided we need. However this approach is, unfortunately, unacceptable since caregivers in one place might be similar and quite different from caregivers in other communities (for a variety of reasons). Thus, doing it this way introduces a bias. The only way to select respondents is to do it in a fully random way. That is, we need to find a way that gives every caregiver a known chance of being selected so that our results (using a subset) are truly representative of the whole (that is, they accurately represent the whole). Given these realities—no lists, need for randomness, caregivers scattered everywhere—we really need to break this question into several sub-questions.

What communities should I go to? Once there, which households should I select? Once in the household, what should I do if, for example, there is more than one possible respondent or if there are no respondents?

These questions lay out one of the most challenging set of issues both conceptually and logistically that that we must deal with in order to conduct a rapid survey. They are the questions of sampling and require detailed, concise, and rigorous methods. Rather than answer all of them here, we would ask you to consult Section II of this manual, where we answer them in some detail. In that section we lay out the issues of sample size and selection processes in detail and provide some alternative ways of answering the question of how to find respondents. Here we have addressed the basis of the challenge at hand. In Section II, we provide answers to these challenges.

Who will carry out the surveys?

The question here also relates to the second phase of conducting a survey, the training of staff to conduct the survey, which is described in more detail below. First we will lay out the different responsibilities in conducting a survey. As we alluded to previously, the question of who is involved in the survey is potentially as broad as all program stakeholders. In terms of setting priorities for a program, assessing its value, and using information to inform decisions, the involvement of a broad group of stakeholders is ideal. However, to keep the actual conduct of a survey manageable, the following represents a minimum of roles/responsibilities that have to be filled:

The Core Team usually consists of a Survey Coordinator, Data Coordinator, and two or three additional members from a combination of the lead organization, and local or national partners, including NGOs, government ministries, the private sector, and others. These primary stakeholders share responsibility for the overall design and conduct of the survey. Members of the Core Team may also function as Supervisors (see below).

The Survey Coordinator is the principal manager in the rapid survey process and should be a staff member of the organization responsible for carrying out the survey with fiscal and managerial decision-making authority. The Survey Coordinator may also be the lead Trainer, if he or she has received the prior survey training.

The Data Coordinator is responsible for either computerized data entry or manual tabulation. This person should have a good background in statistics and computer skills. He or she will supervise the data entry, cleaning, and quality control of data and lead the analysis process.

The Post-Survey Team is responsible for all activities after the field work is completed, including developing an analysis procedure and a plan for follow-up, providing feedback to stakeholders, and writing the survey report.

The Trainer is responsible for training all others (Core Team, Data Coordinator, Supervisors, and Interviewers). This person should have experience in rapid, population-based surveys and should have received some formal training in how to train. This person may not need to design a full training given that such curricula exist (see below), but he or she may need to adapt the training to meet local needs.

Supervisors visit communities prior to the survey to get approval, maps, and/or population information. They supervise the collection of data in the field with strict quality control.

Interviewers pre-test questionnaires. They interview mothers and other respondents in communities. They complete questionnaires, including a review of the documentation. They conduct anthropometric measurements.

Others: Drivers, Logistics Coordinators, Office Administrators, Support Staff, Guides, etc.

The exact number of supervisors and interviewers depends on the size and complexity of the survey. The training materials discussed below provide some guidance on these issues.

How much will the survey cost?

The question of the cost is critical to many of the other decisions related to any survey. In order to answer this question, you have to do some basic logistics planning and budget calculations. Fortunately, useful resources are available for this.

Help is available from the logistics and management planning forms in Annex 2: Calculating the Cost of a Rapid Survey.[14] These forms allow you to develop a survey budget by going through the following steps:

- Step1: Determine the best time of year to undertake the survey taking into consideration local events, such as: holidays; planting or harvest time; rainy season; or requirements of a program planning cycle.
- Step 2: Plan day-to-day activities from survey preparation; interviewing, data analysis, feedback and on to report writing. You can see an example of a plan that allows for a survey to be completed in 28 days below.
- Step 3: Determine what personnel should be recruited for the survey.

[14] Annex 2 is adapted from the KPC Training Materials; you can find the full curriculum for this training—discussed in more detail below—in Annex 1: Key Resources. Another useful resource, the link to which is in Annex 1: Key Resources, is the Malaria Indicator Survey—the module entitled: *Calculating the Cost of the Malaria Indicator Survey.* While developed for a large survey, you might find the budget categories and worksheets helpful.

- Step 4: Make a transportation plan that includes items, such as vehicles, drivers, and fuel.
- Step 5: Plan for editing/ printing and copying
- Step 6: Make a computerized data entry plan
- Step 7: Take the information from steps 1-6 and fill out the following budget worksheets: personnel; transportation; various services and equipment (all found Annex 3).
- Step 8: Use the summary of estimated budget form to calculate the final budget.

The following timeline—running just under 30 days—highlights not only a potential timeline but also demonstrates a number of key activities related to these issues.

Figure M1.4: Potential timeline

DAY 1 Consult with local experts/officials to assess needs, plan survey, form CORE Team. Identify possible field supervisors	**DAY 2** Conduct formative research Identify survey targets & indicators Design questions with stakeholders Recruit supervisors	**DAY 3** Conduct formative research Design questionnaire Develop sampling strategy, analysis plan Recruit field personnel	**DAY 4** Design questionnaire with stakeholders Develop sampling strategy, analysis plan Prepare training materials Recruit field personnel	**DAY 5** Design questionnaire; translate into local language Prepare for training Recruit field personnel Finalize sampling strategy	**DAY 6** Design/translate questionnaire Prepare training, hand tabulation materials Recruit field personnel	**DAY 7** Design/translate questionnaire Recruit field personnel Prepare training materials Select sample areas
DAY 8 Prepare for training: materials, logistics Select sample areas Prepare hand tabulation tables	**DAY 9** Meet w/ supervisors for general training/ overview Conduct pre-test with supervisors; modify questionnaire based on pre-test	**DAY 10** Supervisors visit community leaders to map/identify households for survey Modify/reproduce questionnaires for training	**DAY 11** Prepare for training & field implementation Create data entry template; analysis program Meet with community leaders	**DAY 12** Finalize preparations for training & field implementation Prepare data entry/ analysis programs	**DAY 13** Prepare data entry & analysis programs Begin drafting survey report (sections on objectives, methods, questionnaire)	**DAY 14** Reproduce questionnaires & materials for training workshop Prepare data entry & analysis programs
DAY 15 Train supervisors/ interviewers Prepare data entry/analysis programs	**DAY 16** Train supervisors/ interviewers Practice interviews Prepare data entry program Finalize logistics for field implementation	**DAY 17** Train supervisors/ interviewers Practice interviews Reproduce questionnaires Finalize logistics for field implementation	**DAY 18** Identify survey teams Review protocols Finalize logistics for implementation Identify starting households Finalize entry program	**DAY 19** Data collection Data entry/cleaning Finalize analysis program Finalize hand tabulation tables for workshops	**DAY 20** Data collection Data entry/cleaning Finalize analysis program Finalize hand tabulation tables for workshops	**DAY 21** Data collection Data entry/cleaning Finalize analysis program Finalize hand tabulation tables for workshops
DAY 22 Data collection Data entry/ cleaning Run analysis program Finish tabulation tables	**DAY 23** Hand tabulation with field personnel and other individuals Draft survey report Run analysis program	**DAY 24** Analysis workshop w/ stakeholders & experts Identify health priorities Draft survey report Prepare for feedback sessions	**DAY 25** Finish first draft of survey report Refine action plan Designs ways to display KPC findings Prepare for feedback sessions	**DAY 26** Feedback at community/local level Develop action plan/ M&E plan Plan follow-up research, if necessary	**DAY 27** Brief key stakeholders Plan follow-up research	**DAY 28** Develop action plan and/or M&E plan Plan follow-up research

INCORPORATE LOCAL STAKEHOLDERS AND EXPERTS THROUGHOUT THE PROCESS.

Each survey will have to establish a budget based on its scope and the local context. The logistical costs of a survey in a mountain area of Nepal are far different from an urban survey. The number of indicators, the number of groups to be surveyed (see parallel sampling below), and the time spent in the field will add to the complexity. With more than 420 projects implemented in 61 countries by 55 different international NGOs, the USAID-funded Child Survival and Health Grants Program can provide a benchmark for these costs. Costs as low as US$8,000 have been reported for small surveys, but might not take into account all hidden internal costs for technical assistance and administration. More complex survey designs have run up to US$25,000 and up to US$30,000 in Central Europe. A usual range appears to be in the US$10,000–12,000, but this may vary depending on the region or country where the survey takes place.[15] Rapid surveys cost can be greatly reduced by incorporating community participation in mapping, listing and/or interviewing, or, in the case of LQAS, by incorporating the survey as part of the routine management field visits by project personnel.

TRAINING STAFF TO CONDUCT THE SURVEY

Once you have dealt with all of the preparatory issues including the timing of the survey, you can train the various teams that have a role in the survey. In this section, we summarize the key contents of the training of each group, based on the role of each group. We will summarize the training needs of three main groups: the Core Team, Supervisors and Interviewers, and the Post-Survey Team. While this implies three distinct training sessions, you will see that there is some overlap between the three groups, and you may decide to conduct training in certain topics for everyone together. Once the focus of the survey is determined (purpose, indicators, questions), the training of interviewers is of central importance. We will summarize the key components of each training in this section. Annex 1: Key Resources provides links to training resources that we use in this entire section. These resources were designed to provide training in the KPC survey methodology.

Training the Core Team

As we noted previously, the Core Team usually consists of Survey Coordinator, Data Coordinator, and two or three additional members from a combination of the lead organization, and local or national partners, including NGOs, government ministries, the private sector, and others. These primary stakeholders share responsibility for the overall design and conduct of the survey. Members of the Core Team may also function as Supervisors.

The trainer for the Core Team training should have previous training experience in how to conduct rapid surveys and should be able to adapt existing training materials. The Core Team is directly responsible for designing, organizing, and implementing the rapid health survey. The participants will be responsible for training field staff (supervisors and interviewers and the data entry team) to implement the survey. One of the Core Team members should function as the lead Survey Trainer.

[15] These “benchmarks” are not documented in published format, but come from discussions with many practitioners.

Ideally, some members of the team will have experience with the following:

- Training
- Knowledge of adult education principles
- Experience doing a survey
- Skills in data management and analysis
- Technical knowledge of maternal-child or community health (assuming the survey covers these topics)
- Management skills needed to administer a survey

These participants need technical information about how to design and implement the survey. They need not design the training of supervisors and interviewers from scratch but should be able to adapt already existing materials. This implies that they should be able to access such resources and, in particular, a modular curriculum for training supervisors and interviewers that can be adapted to their learners' needs.

Participants will leave the training with:

- An understanding of the process used and materials needed for implementing a rapid health survey
- A rapid survey design, including a sampling protocol, draft questionnaire, data analysis plan, logistics plan, and budget

You should hold the training in the locality where the rapid survey is going to be conducted immediately prior to starting the survey.

Here is an example of the topical agenda for a five-day training. This and all further examples comes from the KPC Training Guide (see Annex 1: Key Resources for a link to the Training Guide).

Table M1.2: Sample agenda for Core Team training

Day	Learning Session Title
Day 1	
1	Introduction to the Rapid Survey Core Team Training
2	Purpose and Role of the Rapid Surveys
3	Role of the Key Staff in the Rapid Survey Process
4	Identifying Information Needs and Gaps
5	Involving Stakeholders in Rapid Survey Activities
6	Identifying the Target Population for the Rapid Survey
	Daily Evaluation
Day 2	
	Q & A Day 1
7	Overview of Existing Surveys and Tools
8	Adapting the Generic Surveys and Tools
9	Sampling Basics—Why Sample?
10	Sampling Options for Rapid Surveys
11	Bias, Confidence Intervals and Design Effect
	Daily Evaluation

Day	Learning Session Title
Day 3	
	Q & A Day 2
12	Lot Quality Assurance Sampling (LQAS)
13	Selection of Sampling Methodology
14	Community/Household/Informant Selection
	Daily Evaluation
Day 4	NOTE: Sessions 15-17 concern the collection of anthropometric information. Some rapid surveys do NOT include this.
	Q & A Day 3 and Review
15	Purpose of Anthropometry
16	Requirements for Conducting Anthropometric Assessments
17	Anthropometric Data
18	Results Tables Design: Frequencies
19	Results Tables Design: Cross-Tabulation
20	Hand Tabulation
	Daily Evaluation
Day 5	
	Q & A Day 4
21	Quality Control of Data
22	Developing a Data Analysis Plan
23	Finalizing Staffing Decisions
24	Preparations for Training Supervisors and Interviewers
25	Developing a Logistics Plan and Budget

As you can see from the objectives and the sample agenda, this training is less about skill building and more about moving the Core Team through a series of critical design and implementation decisions necessary to conduct the survey and to prepare to train others to do the field work after they have made all key decisions about the content of the survey and the sampling approach.

Training Supervisors/Interviewers

Supervisors and interviewers work closely together to conduct the actual field work of the survey. While supervisors deal with certain logistical issues before the field work (visiting communities prior to the survey to get approval, maps, and/or population information), their main role is supervise the collection of data in the field with strict quality control. The people they supervise are the interviewers.

Interviewers also start their work before the actual interviewing in the field by pre-testing questionnaires and ensuring that all materials are clear for everyone to use. In the field, they interview mothers and other respondents in communities. They complete questionnaires and may, if appropriate, conduct anthropometric measurements. The training of these two groups is critical to the success of the survey since they are directly responsible for gathering the information.

The team of supervisors and interviewers normally consists of 12 to 18 people. One supervisor and two interviewers usually form an Interview Team, although the ratio of supervisor to interviewer may vary. Supervisors may also be members of the Core Team.

Ideally, the individuals identified as supervisors and interviewers should have:

- Experience participating in other surveys
- Local language skills
- Knowledge of maternal-child health or community health

The focus of this training workshop is to ensure that supervisors and interviewers understand all of the questions on the rapid survey questionnaire and the procedures for selecting both households and respondents. In addition, this training workshop provides an opportunity for interviewers to practice interviewing skills and for supervisors to practice supervision skills.

Participants will leave the training with:

- An understanding of how the health data collected relate to the objectives of the program and the rapid survey
- All materials and information about logistics and timeframes for the rapid survey interviews
- A clear understanding of each question on the rapid survey questionnaire, the indicator each question is designed to measure, and why that particular indicator is considered important to the program
- Experience, through practice using the final rapid survey questionnaire
- Skills in supervision, interviewing, and good survey techniques

The Core Team should conduct this training workshop and carry out a field test of the adapted rapid survey questionnaire. This training workshop should occur immediately **after** the training of the Core Team and immediately **before** conducting the actual field survey. Ideally, the training workshop would be held at a training center near the location where you will conduct the survey. The advantages of being near the survey area are that 1) field tests can then be conducted in the appropriate language, and 2) the local events calendars and lexicon will more closely reflect the situation expected while conducting the survey[16].

Here is an example of the topical agenda for a four-day training.

Table M1.3: Sample agenda for Supervisors/Interviewers training

Day	Learning Session Title
Day 1	
1	Introduction to the Training
2	The Purpose and Role of the Rapid Survey
3	Role of the Core Team, Supervisors and Interviewers
4	Reviewing the Rapid Survey Questionnaire
5	Selection of Households and Respondents
	Daily Evaluation
Day 2	
	Question and Answer (about the previous day's work)
6	Proper Interviewing Techniques
7	Importance of Informed Consent and Confidentiality
8	Using Documentation (things like Local Events Calendars to assess a child's age)
9	Giving Feedback on Practice Interviewing

[16] The training materials referenced here define clearly the meaning of "local events calendar" and "lexicon."

Day	Learning Session Title
	Daily Evaluation
Day 3	NOTE: Sessions 10 and 11 concern the collection of anthropometric information. Some rapid surveys do NOT include this.
	Question and Answer (about the previous day's work)
10	Measuring Weight
11	Conducting Standardization Testing
12	Practicing Interviews
	Daily Evaluation
Day 4	
	Question & Answer (about the previous day's work)
13	Field Test and Revise the Rapid Survey Instruments
	Evaluation and Closing

Notice that in contrast to the training of the Core Team, this training is really about skill building—preparing people to do things like properly select households and respondents and use the survey questionnaire and all materials necessary to conduct the survey.

Training the Post-Survey Team

The Post-Survey Team training is designed for individuals who will provide leadership in the process of analyzing the data from the rapid survey and who will make management decisions based on that data. Even though these activities occur after you have conducted the field work for the survey, we place them here since they concern training.

The Post-Survey Team normally consists of the Core Team (four to six people) and other key staff from the local partners and other stakeholders. The Post-Survey Team will be directly responsible for analyzing and using data from the rapid survey. In addition, the training will help this team involve other stakeholders in the analysis process. Ideally, the individuals identified for the Post-Survey Team will have had previous experience in:

- Conducting a survey
- Data management and analysis
- Training
- Also desirable: technical knowledge of maternal-child or community health

The trainer for the Post-Survey Team training should have previous training experience in how to conduct rapid surveys and should be able to adapt existing training materials but need not create his or her own training design. He or she is responsible for using this module to train the Core Team and other participants to begin the analysis process and to outline how to involve other stakeholders in the process. The rapid health survey is a population-based survey that is statistically valid and focuses on critical health indicators. To ensure that the rapid health survey is of maximum value, it is essential that the Core Team be trained in how to use the results in project management and how to share the results for further analysis.

Participants will leave this training with:

- An understanding of what to look for when tabulating/analyzing rapid survey data
- A preliminary exploration of frequencies and differentials in a rapid health survey dataset
- A comparison of rapid survey findings with results from other data sources

- A set of reasonable intermediate and final targets based on rapid survey baseline data or use the rapid survey data to assess achievement of targets and determine whether differences are statistically significant
- Rapid survey data to identify health problems and possible intervention activities/strategies and the level of effort needed for each intervention
- Decisions on follow-up studies/activities that will be conducted
- A list of other levels (e.g., health-facility level) where change must occur in order to effect changes at the population level, and decide whether the program should undertake studies at these levels as well
- A draft rapid survey report
- A plan for discussing rapid survey data with program communities and other stakeholders

You should conduct this training immediately after the rapid survey has been completed and data have been tabulated either manually or with computer software.

Here is an example of the topical agenda for a two-day training.

Table M1.4: Sample agenda for Post-Survey Team training

Day	Learning Session Title
Day 1	
1	Introduction to the Post-Survey Analysis Team Training
2	Making Decisions Using Rapid Survey Data
3	Comparing Findings with Other Surveys and Data Sources
4a	(For Baseline Surveys) Using Rapid Survey Baseline Results to Establish Levels of Effort and Targets
	OR
4b	(For Mid-term or Final Surveys) Assessing Achievement of Targets
	Daily Evaluation
Day 2	
	Q&A: Day 1
5	Identifying Follow-up Activities
6	Writing the Rapid Survey Report
7	Discussing Rapid Survey Data to Community Members and Other Stakeholders

CARRYING OUT THE FIELD WORK

The field work phase involves conducting interviews in selected communities. **Quality control is critical to the data collection process.** The purpose of quality control procedures is to maximize the performance of the interviewers and get the best possible data, given the circumstances of the local context. Another way to put this is to say that the purpose is to reduce non-sampling errors (we will examine the issue of sampling errors below).

Before sending survey teams into communities, the Core Team should have a clear strategy for maintaining quality throughout the data collection process. The team should focus on this process in the training of supervisors and interviewers. It helps to create a field implementation checklist[17]" such as the example in Figure M1.5, so that each survey team can take a daily inventory of all supplies and equipment before going into the field.

Supervisors should not conduct interviews. Their primary role is to support the interviewers, serving as the first point of contact when interviewers encounter problems in the field and assessing and maintaining the quality of data collection. Supervisors should observe every interviewer they are responsible for conduct at least one interview every day and complete a quality improvement checklist (see the example in Figure M1.6) for every interview they observe. The purpose of this checklist is to provide a framework for giving immediate feedback to interviewers on their work so they can improve it as they go.

Supervisors will not have the chance to observe every interview conducted by interviewers in their teams, but they are responsible for reviewing every questionnaire for errors. They should do this while in the field, so that any problems can be resolved immediately. The supervisor should indicate any changes or notes on the questionnaire using a colored pen or pencil. Once a supervisor has finished reviewing a completed questionnaire, he or she should sign or initial the last page to indicate that the questionnaire has been checked for quality.

Members of the Core Team should visit the field periodically to assess data collection activities as well. They can also observe interviews and use the same quality improvement checklist to provide feedback to interviewers.

Before sending interviewers out into the community, it is helpful to have a meeting with all interviewers and supervisors to do the following:

- Last-minute troubleshooting
- Confirm availability of all necessary supplies (see the example in Figure M1.5)
- Confirm assigned locations of each survey team for that day
- Review community entry protocol (visit local leaders, health workers)
- Review household selection protocol
- Review respondent selection protocol

All interviewers and supervisors should meet daily as a group to share experiences and problems. At a minimum, each survey team should meet at the end of each day to submit completed questionnaires to the supervisor and discuss any problems or receive follow-up training. It is also helpful if supervisors meet as a group with the Core Team for daily follow-up.

You can create a daily reporting form for recording field problems. Examples of problems that should be documented as they arise and submitted to the Core Team daily include the following:

- Problems with household or respondent selection
- Problems with the questionnaire, for example:
    - List of questions not understood by mothers

[17] Please note that we have not yet introduced some of the terms used in this tool and the quality improvement checklist. We introduce some further on in this handbook, but some are beyond its scope. You can find information about all of them in the training and resource materials described in Annex 1: Key Resources.

    - Incorrect skip patterns
    - Inappropriate terminology or wording
- Non-functioning or lost equipment
- Other problems encountered by interviewers

Figure M1.5: Example of a field implementation checklist

(to be completed by each survey team)

Transportation
- ☐ Car/van
- ☐ Driver
- ☐ Petrol
- ☐ Community guide (someone who is familiar with the communities—can also be the driver)

Depending on resources and the layout of the project area, more than one survey team may be assigned to the same vehicle and driver. The team(s) and the driver should agree upon the drop-off and pick-up times and locations within each sample area.

Food and Other Provisions
- ☐ Drinking Water
- ☐ Bag lunch/food allowance
- ☐ First Aid Kit

Survey Equipment

For interviewers
- ☐ Pencils/pens/erasers
- ☐ Clipboards
- ☐ Adequate copies of the questionnaire (for at least one day's worth of interviews)
- ☐ Medicines for display during the interview, for example (depending on the survey questions):
  - ORS packet
  - Vitamin A capsules
  - Iron/folate tablets
- ☐ Tools for random selection, such as:
  - Empty bottle or other designated object (if using spin-the-bottle technique)
  - Coin (for flip-the-coin technique)
  - Random number tables
  - Blank paper
- ☐ Quick reference sheet with protocols for household and respondent selection
- ☐ Necessary equipment for anthropometric measurement, for example (if anthropometry is included):
  - Scales
  - Measuring boards
  - Tape measures or MUAC insertion tapes (for measuring mid-upper arm circumference)

For Supervisors
- ☐ List of selected communities and number of clusters/interview sites in each; each survey team and their cluster/interview site assignments (particularly important if more than one team will be conducting interviews in the same community)
- ☐ Extra copies of questionnaires
- ☐ Extra pens/pencils/erasers
- ☐ Extra medicines to display (e.g. vitamin A capsules, ORS packets, iron/folate tablets)
- ☐ Quality-control checklists
- ☐ Maps/listing of households in the sample area

Figure M1.6: Example of a quality improvement checklist

Interviewer: ______________________ **Date:** ______________________
Supervisor: ______________________ **Community:** ______________________

Observe and evaluate a minimum of one (1) interview conducted by each Interviewer in your Interview Team each day. Use this form as you observe the Interviewers. While you are observing, do not talk with the Interviewer being evaluated. Completely fill in this form. When the Interviewer is finished conducting the interview, review this form with the Interviewer in private. Quickly discuss each point, pointing out both the Interviewer's strong points and the areas where improvement is needed. Remember: The purpose of this form is to *document* the quality of the interviews, encourage the Interviewer to continue doing what she/he is doing correctly, and to *improve* the Interviewers' performance.

Interview start time: ______________________

	DID THE INTERVIEWER. . .	NO	YES
1.	Select the *household* correctly?		
2.	Select the *respondent* correctly?		
3.	If the intended respondent was not at home, did the Interviewer use the proper protocol (e.g., find the respondent if less than ____ minutes away)?		
4.	Introduce him/herself correctly?		
5.	Read the consent statement at the beginning of the interview and get permission without coercion?		
6.	Correctly record information on the cover page (such as interview date, name of community, mother's/child's name, mother's/child's age/date of birth, child's sex)?		
7.	Correctly calculate the child's age in months from the DOB?		
8.	Use the events calendar properly, if it was needed?		
9.	Speak clearly during the interview?		
10.	Use culturally appropriate body language?		
11.	Have neutral facial expressions/body language (did not react positively or negatively to the respondent's answers)?		
12.	Refrain from asking leading questions that might have influenced the respondent's answers?		
13.	Read the questions exactly as they were written?		
14.	Repeat the questions exactly as worded when the respondent gave a response that was not very clear? Use probes when the response still was not very clear?		
15.	Write legibly on the questionnaire?		
16.	Follow the skip patterns correctly?		
17.	Read responses aloud when he/she was supposed to?		
18.	Prompt the mother for all answers (say "Anything else?") for questions that allow multiple responses?		
19.	Weigh/measure the child correctly?		
20.	Thank the respondent for the time spent and involvement in the survey?		
On a scale of 1 (needs more training) to 10 (excellent), I rate the interviewer's performance during this interview as follows (circle one): Needs More Training 1 2 3 4 5 6 7 8 9 10 Excellent			

Interview end time:________ Approximate Duration of Interview:__________________ minutes
Other Comments/Plan of Action for Making Improvements: ______________________________________

USING SURVEY RESULTS

Once all the data have been collected and initial field-based corrections have been done, the survey team can tabulate, analyze, and discuss findings with stakeholders. Once again, in this section we do not present all the details for this process but lay out some general issues to consider while referring the reader to more detailed resources in the KPC training materials in Annex 1: Key Resources. In addition, Section II (especially the section on LQAS) provides more detail on how you can use data to prioritize support to field staff implementing the program.

Keep in mind that we do not wait until data has been collected to think about data entry, tabulation, and analysis. As we have seen, the analysis plan is actually part of survey preparation and is finalized during the Core Team training. Thus, at this stage, the team is merely putting into practice the plan it has already developed.

TABULATING/ENTERING DATA

While computers have become the easiest way to enter and check data it may be useful to hold a workshop in which key stakeholders can participate in "hand tabulation" of data. For staff and stakeholders with less experience in rapid surveys, such a workshop is a way to demystify how we move from questionnaires to the results we analyze. In this section, we discuss only hand tabulation. For computerized data entry and checking, please see the resources from UCLA Department of Epidemiology in Annex 1: Key Resources.

Manual (hand) tabulation gives a hands-on feeling for what the data mean to a larger number of people. If local partners and stakeholders opted not to be directly involved in data collection activities, invite them again to participate in a hand tabulation workshop. This is a prime opportunity for all stakeholders to work directly with the data and identify and prioritize problems as a group. By being transparent in terms of how the data are collected, analyzed, and interpreted, a project can use a hand tabulation workshop to build consensus among stakeholders. Hand tabulation is also a good way to validate results generated by a computer.

All Core Team members should participate in a hand tabulation workshop and can facilitate the tabulation process. However, other individuals should have the opportunity to work directly with the data. Invite other individuals who either have a stake in the project or are affiliated with agencies and institutions that work in the same geographic area. Supervisors and interviewers can help tabulate. You do not have to limit participants to individuals who work in the health field only. For example, you can also invite local communication specialists, water and sanitation experts, or qualitative researchers. In planning the workshop, keep in mind that a group that is too large will be hard to facilitate. Aim to have a group that is manageable—given time, space, and other constraints—yet includes people who represent different perspectives on child health and survival. With a diverse group of workshop participants, your project might gain insight into why certain problems exist and how those problems can be addressed.

You do not need to tabulate every question from the survey. In addition, if the Core Team has decided to conduct the data analysis by computer, you do not need to hand tabulate all of the key program indicators during the workshop. The objective of a hand tabulation workshop is to ensure that other individuals understand the data, see value in them, and ultimately use them to prioritize problems and develop solutions.

To tabulate questionnaires, you will need the following:[18]

- Completed and corrected rapid survey questionnaires
- List of rapid survey indicators that you plan to tabulate together
- Tabulation and analysis plan
- Photocopies of tabulation tables for workshop participants
- Clear instructions for workshop participants on how to tabulate indicators by hand (which survey question responses are necessary to construct which indicators).

Once everyone is in place, divide the participants into groups to work on perhaps a single indicator. During the workshop, members of the Core Team should review each group's output and provide feedback. After hand tabulating indicators in small groups, everyone can reconvene as one large group to share results based on their tabulations. At this point, the group can begin analyzing and discussing findings from the survey. It might even be helpful to explore subgroup differences. For example, are certain indicators different for boys versus girls? for children of young women versus those of older women? for children whose birth was inadequately spaced versus those whose birth was adequately spaced?

The LQAS Training Manual a link to which is in Annex 1: Key Resources provides examples of hand tabulation tables for that approach. Hand tabulation is merely a matter of using a tally sheet to enter counts for the numerator and denominator for each indicator. Thus, for an indicator such as, "Percentage of mothers with children age 0–23 months who report that they wash their hands with soap/ash before food preparation, before feeding children, after defecation, and after attending to a child who has defecated," you would tally all the mothers of children 0–23 months of age, basically counting all surveys for which respondents answered the questions about hand washing. Then you would count the surveys for which respondents answer yes to all of the questions concerning hand washing 1) before food preparation, 2) before feeding children, 3) after defecation, and 4) after attending to a child who has defecated.

Note that the group should only tally as "correct" questionnaires on which respondents said yes to all four questions about hand washing. While there may be other tabulations about individual behaviors related to hand washing, the indicator of interest here concerns correctly practicing all four behaviors. Thus a questionnaire should receive a tally mark only if the respondent reported practicing all four of them.

[18] Important note: Before you begin tabulating data, you must first fix any errors that were made during data collection. Data cleaning—as this is called—involves identifying and correcting those mistakes. As we noted in the previous section on field work, error checking can take place during field work when it is still possible to correct mistakes. As interviewers complete interviews, supervisors should review the completed questionnaires to make sure that interviewers filled them out correctly. Supervisors can then follow up with interviewers to correct any mistakes identified, revisit respondents (if necessary), and make sure that those mistakes are not repeated in the remaining interviews. If there are good quality control procedures during data collection, then the error-checking process should not be difficult.

ANALYZING DATA

Again, we could write at length about data analysis, and your exact analysis approach depends on your specific needs and how you have collected the data. In Module 3 on LQAS, we discuss ways to think about analyzing results by "supervision area" (SA). This section provides introductory principles and thoughts on the process.

Most analyses are limited to frequencies (counts) of each survey question. Although a rapid survey is a (relatively) small-sample survey, we encourage you to do more in-depth problem analysis using the data. For example, simple cross-tabulations might highlight differences that exist between groups in the target population. These differences might warrant further attention, for example, through qualitative research.

Here is a simple example of what we mean by a cross-tabulation. The ones you choose are a function of your interests, objectives, and decision making needs.

Table M1.5: Cross-tabulation example

	Use a Family Planning Method	**Do Not Use a Family Planning Method**	
Women under 25 years of age	Women under 25 using a family planning method	Women under 25 not using a family planning method	All women under 25 years of age
Women 25 years of age and older	Women 25 and over using a family planning method	Women 25 and over not using a family planning method	All women 25 years of age and older
	All women using a family planning method	All women not using a family planning method	Total number of women responding

In this example, we use a cross-tabulation to compare family planning practices by the age of women in the survey—comparing younger women with older women to see whether a woman's age is associated with use of family planning methods (is there some evidence of an association between being younger and using family planning, for example).

Your data analysis process is also a function of where the survey occurs in the life of a program. If the survey you are analyzing is at baseline, you are likely using the results to set priorities and you are likely to spend more time exploring various connections (via cross-tabulations, for example) than you might at other times. For analysis of mid-term or end-term evaluations, you should have a much clearer sense of the issues you must address and comparisons you want to make based on objectives you have set.

In general in your analysis you are considering the following issues, whether you are looking at basic frequencies or cross-tabulations:

- Do the results confirm what you and stakeholders expected? Which results are surprising or troubling?
- How do these results compare with data from other sources or with targets you have set?
- What gaps between current practices and coverage concern you in that they need to be addressed or do not conform to targets you set?

Just as we encouraged you to hold a workshop to tabulate data, we would encourage a process for conducting the analysis with a group of stakeholders. Some or all of this could be done at a tabulation workshop if you provide some computer (or previously prepared hand) tabulations and combine them with the hand tabulations done in the workshop. The key is to walk through all indicators of interest to the project and invite participants (in a combination of small and large groups) to discuss results and possible reasons for those results. You should also pause during analysis to ask what other information you may need to collect using other methods such as health facility assessments or in-depth interviews with caregivers. Further discussion and examples of analysis of rapid survey data is available online at the Software Training Manual from UCLA. The link to this excellent resource is found in Annex 1: Key Resources.

PRESENTING RESULTS

Once you have conducted the analysis you will want to present the results in various forms—via verbal presentations and in written reports. The following provides a sample format for a typical written report. It is adapted from the KPC training materials referenced in Annex 1: Key Resources, and the training materials provide greater detail on the contents of each section.

Rapid Survey Report
Table of Contents

I. Executive Summary

II. Background

A. Project location and background of the area
B. Characteristics of the target beneficiary population
C. Social, economic and health conditions within the project area
D. National standards/policies regarding health issues of interest
E. Overview of the project: goals, objectives and activities
F. Results of qualitative studies
G. Objectives of the rapid survey

III. Partnership Building

A. Methods of engaging local partners/stakeholders in the rapid survey
B. Specific roles of local partners/stakeholders in the rapid survey

IV. METHODS

A. Questionnaire development
B. Indicators
C. Sampling design
D. Training approach
E. Data collection and quality control procedures
F. Data management/data analysis

V. Results

Tables of results and graphics for principal findings

VI. Discussion

A. Discussion of key findings from the KPC and programmatic implications
B. Next steps in information gathering
C. Action Plan for community feedback and dissemination of findings

VII. ANNEXES

Annex A: Map of Project Area with clusters/interview locations identified
Annex B: Logistical Preparations and Schedule
Annex C: Survey Questionnaire in English and [local language]
Annex D: Sampling Frame
Annex E: Training Guide and Schedule for KPC Survey Training
Annex F: Manual Tabulation Tables
Annex G: Computer Tables for Each Question
Annex H: Breakdown of Costs for the Rapid Survey

Example Frequency Table: Percentage of children age 12–23 months appropriately vaccinated before the first birthday:

n= 261

	Number	Percent	Cumulative Percent
Received all vaccines	141	54%	54%
Did not receive all vaccines	120	46%	100%

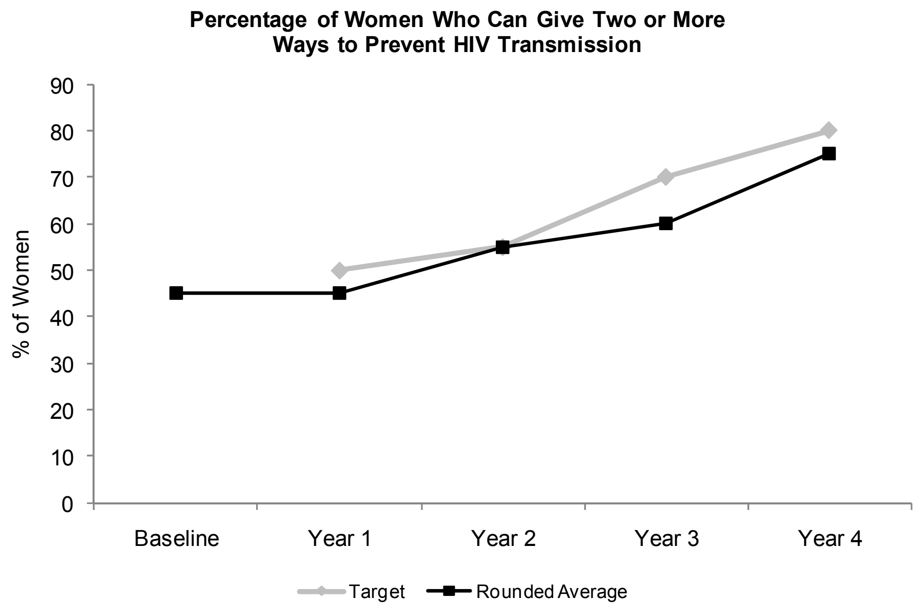
Example Cross Tabulation: Sick children age 0–23 months who received increased fluids and continued feeding during an illness in the past two (2) weeks by mother's age. For cross-tabulations it is important to decide whether to include row or column percentages. Here, since we want to compare the behavior of mothers of different ages we provide column percentages. So for the first cell we have (40/59)*100=67.8%

n = 119

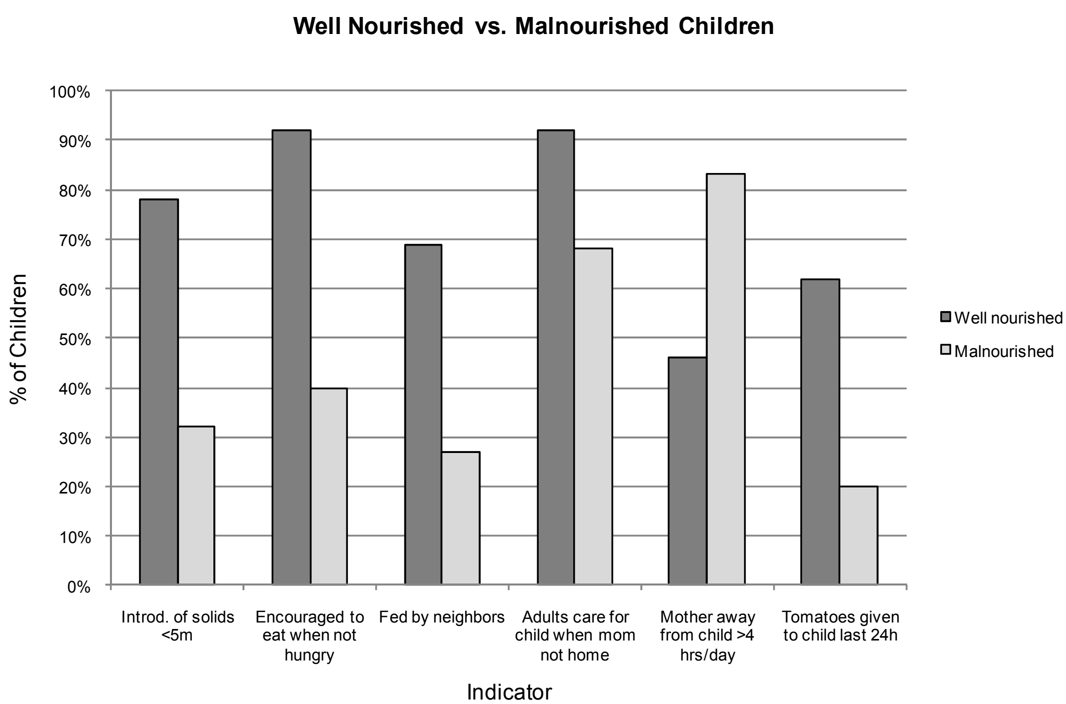
	Mother's Age < 25 years	Mother's Age ≥ 25 years	
Received both more fluids and the same and/or more food	40 (67.8%)	7 (11.7%)	47 (39.5%)
Did not receive both more fluids and the same and/or more food	19 (32.2%)	53 (88.3%)	72 (60.5%)
Totals	59	60	119

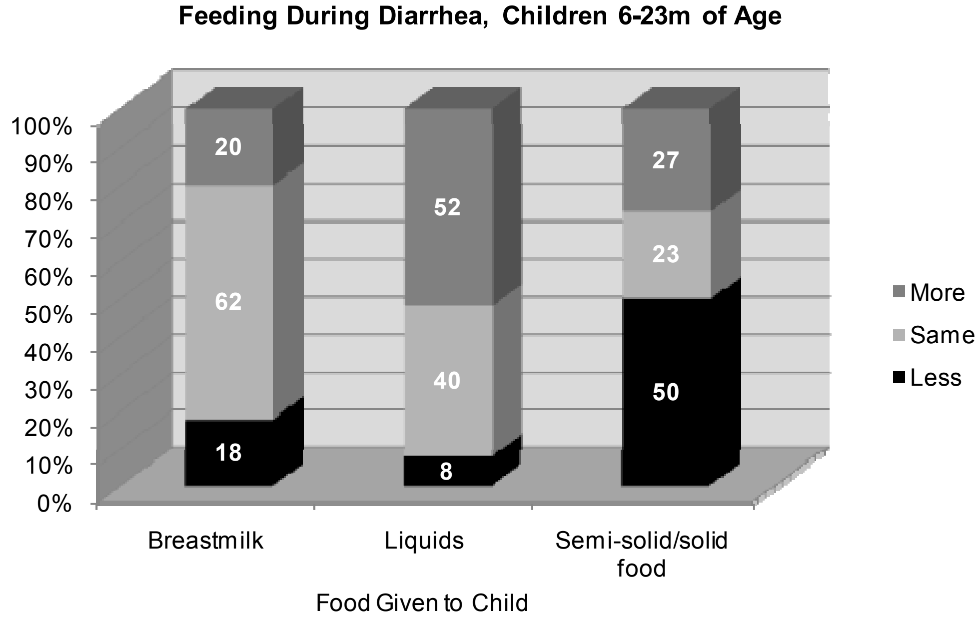
Carefully selected graphics can be useful to display indicator data. Be sure to label them clearly.

- Line graphs can clearly show change over time:

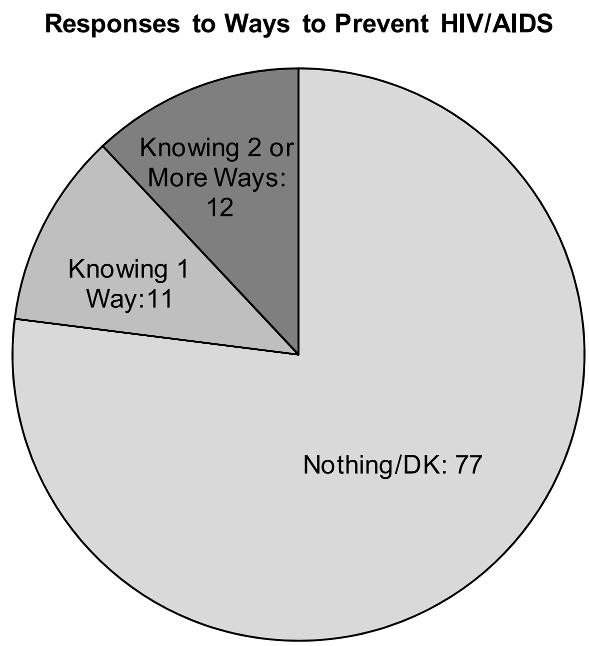


- Bar graphs are especially good at showing comparisons of two variables—cross-tabulations, baseline vs. final, etc:





- Pie charts can be used effectively to show the breakdown by response categories:



The appropriate way to present data is another area that goes beyond the scope of this document. However, these few examples represent key ways that we often show data using graphics and we would encourage you to use these formats while keeping the amount of information you share limited and the graphs simple.

Finally, we would encourage you to consider how to prevent key findings with community members in areas where you are carrying out your program. The format for doing so depends on your local situation and probably involves face-to-face meetings. The purpose of these meetings is to give feedback to communities, to promote analysis of problems at the community level, to encourage commitments from stakeholders regarding involvement in program activities and to hear back from community members about changes and or challenges they are observing in regards to program activities and outcomes.

SECTION II: SAMPLING

INTRODUCTION

In Section I, we introduced the purpose of rapid surveys and outlined the process for carrying them out. In Module 1, on the steps for carrying out rapid surveys, we asked two related questions:

1. Whom should I survey?
2. How do I find respondents?

These two questions are both related to the issue of sampling—the process of using a subset of a population in order to make generalizations or draw meaningful conclusions about the whole population. In most health programs, the population of interest is found within a specific program area and consists of people targeted by the program. In many of the examples we have used so far, the beneficiary population has been children under five years of age and women of reproductive age, although for the purpose of the considered health surveys we have targeted caregivers of children 0–23 months. A given health program might focus on other populations, but the point is that we use rapid surveys in order to learn something about that population, usually to set coverage targets for the population or assess whether coverage has reached that targets already set for the population. We noted that in conducting rapid surveys, one of the most challenging issues concerns sampling, and the remainder of this handbook will focus on those issues.

This section introduces key sampling terminology used in Modules 2 and 3 to describe in detail two sampling methods: two-stage cluster sampling and LQAS.

SAMPLING CONCEPTS AND STRATEGIES

Suppose that we are managing a maternal and child health program that operates in District X. We are in the process of re-organizing activities, so we need information on women and children in the district. It would be too costly and time consuming to do a census. We have information from health centers, but we know that many people do not use them. The only way to find out what is happening with women and children even if they do not all use health centers equally and in a cost-effective manner, is to take a sample. Sampling is the process of selecting units (e.g., people, organizations) from a population of interest so that by studying the sample we can generalize our results back to the population from which they were chosen.[19]

There are many different ways that we could take a sample, but we want a method that allows us to make an estimate that is close to the true level in the population but that does not overburden us with data collection. Methods that allow us to make these estimations are based on scientific probability sampling, which also allow us to calculate the sampling error (the effect of interviewing a portion instead of the whole universe of interest or, more simply, the difference between our estimate and the true level in the population). A probability sample is defined as one in which the units are selected randomly with known and non-zero probabilities. The term excludes purposive sampling, quota sampling, convenience sampling, and other uncontrolled

[19] Trochim, William M. The Research Methods Knowledge Base, 2nd Edition. http://www.socialresearchmethods.net/kb/, (October 20, 2006).

non-probability methods, since they cannot provide precision and/or confident evaluation of survey findings.

This section is divided into 4 parts: (I) preparation for choosing a sampling method; (II) basic sampling terminology; (III) brief descriptions of five sampling methodologies: simple random sampling (SRS); systematic sampling; stratified sampling; cluster; and LQAS (the final two methodologies are developed fully in Modules 2 and 3); and (IV) additional sampling topics

PART I: PREPARATION FOR CHOOSING A SAMPLING METHOD

Before deciding on a sampling method, it is important to: (1) take stock of what information you need, (2) understand the reality of where the survey will be implemented, and (3) understand the constraints of what time and budget are available for the survey.

Consider as an example a program addressing maternal, newborn, and child health, as well as HIV/AIDS. Suppose the program area is located mostly in a mountainous region with difficult access to many villages, which are often accessible only by narrow dirt roads. Information must be collected during the month of September, before the planning cycle starts in October, but after the August festivities. Indicators have been developed for each intervention area.

1) Take stock of what information you need.

Before deciding on a sampling methodology, it is important to analyze the type of information that is proposed for the survey and to make adjustments to this information in order to simplify sampling. The following questions will help you with this analysis.

- From what group or groups do you need information?
- Could you re-write the indicators to reduce the number of groups you will need to sample?
- Is the event being studied rare or common?
- Could you re-write the indicator to one that closely measures the same thing but is not so rare?
- Would it be better to obtain this information through another methodology (e.g., qualitative data collection, health facility survey), rather than a population-based survey?

Table 2.1 provides examples of indicators that a program manager might propose and a way of organizing the analysis of this information.

Table 2.1: Indicator examples

Information/Indicators	Sampling group	Is the event rare or common?	Comment
% of children 0–5 months of age who are exclusively breastfed	Children 0–5 months	Common	The event being studied is type of feeding of children 0–5 months who should be exclusively breastfed.
% of children 0–23 months whose births were attended by skilled personnel	Children 0–23 months	Common	The event being studied is type of attendance at birth. Of course, all of these children were born.
% of mothers who were vaccinated with TT before the birth of her youngest child	Mothers	Common	The event being studied is TT vaccination status of mothers
% of pregnant women who take iron folate	Women who are pregnant at the time of the survey	Relatively rare	The event being studied is use of iron folate in women who are pregnant. There are relatively few pregnant women at the time of any survey.
% of adolescent males (15–24 years) who know three ways of preventing infection with HIV	Adolescent males (15–24 years)	Common	The event being studied is knowledge of HIV prevention by adolescent males. Adolescent males are common in a population.
% of pregnant women with preeclampsia who are referred to an appropriate facility	Women who are pregnant and have preeclampsia	Rare	The event being studied is the response to pregnant women who have preeclampsia. There are relatively few pregnant women at the time of any survey, and only a small percentage of them will develop preeclampsia.

In this example you can see that there are six different sampling groups:

- Children 0–5 months
- Children 0–23 months
- Mothers
- Pregnant women
- Adolescent males (15–24 years)
- Pregnant women with preeclampsia

There are several problems with these sampling groups. First, having this many sampling groups can complicate sampling. Second, one group (mothers) is much larger than the other groups, which adds another level of complication. Third, two indicators correspond to rare events: pregnant women (relatively rare) and pregnant women with preeclampsia (very rare). Designing a survey to measure information about these rare events would require a very large sample and visiting many households to find the appropriate respondents.

Information collected through a rapid population-based survey should be simplified so that the survey can be as efficient as possible, while maintaining rigor. We can simplify the example above, while still providing important information. First, reword the indicator for pregnant women so that it can be measured from one of the other sampling groups (children 0–23 months). Second, narrow the definition of the indicator for mothers so that it is consistent with one of the other sampling groups (children 0–23 months). Third, eliminate information about the response to preeclampsia from the survey. Other techniques can be used to collect information about rare events; however this is outside of the scope of this guide. The table below shows the result of making these changes (*underlined in italics*).

Table 2.2: Modified indicator examples

Revised indicators	Sampling group	Comment
% of children 0–5 months of age who are exclusively breastfed	*Mothers of* children 0–5 months	To obtain this information, mothers of children 0–5 months are interviewed.
% of children 0–23 months whose births were attended by skilled personnel	*Mothers of* children 0–23 months	To obtain this information, mothers of children 0–23 months are interviewed.
% of mothers *of children 0–23 months* who were vaccinated with TT before the birth of her youngest child	Mothers *of children 0–23 months*	To obtain this information, mothers of children 0–23 months are interviewed.
% of *mothers of children 0–23 months who took* iron folate *while they were pregnant with the youngest child*	*Mothers of children 0–23 months*	To obtain this information, mothers of children 0–23 months are interviewed.
% of adolescent males (15–24 years) who know 3 ways of preventing infection with HIV	Adolescent males (15–24 years)	This information is from males 15–24 years.
Eliminate: % of pregnant women with preeclampsia who are referred to an appropriate facility	*Eliminate*: Women who are pregnant and have preeclampsia	Find another method to collect this information.

With these changes, the number of sampling groups has been reduced to three: mothers of children 0–5 months, mothers of children 0–23 months, and adolescent males (15–24 years). Now the survey will be manageable.

2) Understand the reality of where the survey will be implemented.

As you plan the study, think about the challenges of where you will implement the survey. The following questions are useful:

- What is the terrain like?
- How difficult is it to move around?
- How densely populated is the project area?

In this example, the survey will be carried out in a mountainous area with many villages that are isolated and often accessible only by narrow dirt roads. It will be logistically challenging and expensive to reach these villages. You still need information from the population, so you will make the effort to visit some of these villages; however, it is important to choose a sampling method that allows you to visit fewer sites. For example, you might not want to use simple random sampling, where perhaps you would have to visit 300 sites (one site for each interview). Instead, you might choose methods such as cluster sampling or LQAS where you visit fewer sites (i.e., 30 for 30-cluster and 95 for LQAS). However, if your program is in a densely populated urban area and you have a list of all participants; it might not be difficult to use SRS to choose 300 respondents and to visit each one. Therefore, you would theoretically at least have the possibility of choosing SRS. It does happen, for example for surveys conducted in a food distribution program area. These programs usually have a list of all the households in the area. SRS could easily be performed on this list in order to find names of respondents to be interviewed.

Does a household list exist for every house in the project area? In most areas, household lists are not available. Although it could be possible to include a preliminary step of creating a household list for the entire project area, this usually is too time-consuming and expensive. Methods, such as cluster sampling and LQAS, can be implemented without the requirement of household lists for the entire project area, so are appropriate choices in those situations. There are some instances in which household lists already exist. We already mentioned food distribution programs. Other programs include a census as part of program implementation, and the household list generated by the census can be used for the survey.

3) Understand the constraints of time and budget available for the survey.

You should choose a sampling method that allows you to collect information so it can be used in a timely manner for decision-making. In the above example, the study must be completed before the October planning cycle but after the August festivities. Therefore, choose a sampling methodology that allows you to collect the information in one month. The two methodologies presented in detail later in this guide would be appropriate choices.

Usually, there are limited resources for conducting the survey. It is important to adjust the survey design to meet this reality.[20] You can take steps to contain costs. The following are examples:

- Choosing a sampling methodology where you visit a smaller number of sites will reduce costs in situations where there are logistical challenges in reaching sites. As noted, the two-stage cluster sampling approach is one such methodology, as is LQAS.
- Accepting a lower level of precision allows you to collect a smaller sample. For program management purposes, 10% is a common level. This means that if a coverage level of 45% is obtained from the survey, the actual level could be between 35% and 55%.[21] Precision could be increased to 5%, which means that the actual level would be between 40% and 50%. As always, increased precision requires an increased sample size with a corresponding increased cost. Deciding on the precision level should balance cost and time constraints with what is needed to provide useful and meaningful information.
- Using household or individual lists already compiled or choosing a sampling methodology that does not require these lists for the entire program area. These options reduce the time spent on creating this list.

[20] This should not stop practitioners from advocating within their organizations and with donors for the dedication of more resources to these cost effective ways of obtaining real population-level data.

[21] Generally when an indicator is at a low level (i.e. below 30%), an absolute margin of error of up to 10% becomes quite important, and proper statistic principles would reject establishing estimates for such low indicators. Estimating that an indicator is at 20%, with a confidence interval between 11% and 29% shows evidently very poor precision. However for the manager at a local level, this information is still useful in showing very poor performance, and determining that within the next two years, for example, the indicator needs to be brought to 50% throughout the district. While imprecise, the estimate from the relatively small sample of the rapid survey remains useful.

In the end, most programs using rapid survey approaches choose sampling methods and sample sizes that reduce costs and time yet allow for enough precision for management purposes. This is why we introduce two stage, 30-cluster sampling with total sample sizes of 300 respondents (10 respondents per cluster) and LQAS samples of 95 respondents. In general (and given some assumptions we develop more below) these approaches provide precision levels of 10% or better (in absolute terms as described above) for all indicators of interest, while minimizing the number of locations to be visited by survey teams. We will however discuss when and how this sample size can be increased.

In addition, in all surveys (small or large), non-sampling errors (coverage errors, errors committed in survey implementation, interviewer errors, and data processing, etc.) are usually the most important sources of error. It is therefore important to minimize them in any type of survey implementation. In Module 1, we discussed ways to do this in part through better supervision during field implementation. Additionally, in order to facilitate accurate implementation of the survey, the sampling design should be as simple and straightforward as possible. Again, the two-stage cluster and LQAS sampling methodologies are relatively easy to implement and thus we can better maintain quality.

While we present these specific approaches and sample sizes and strongly encourage you to consider using them, we realize that some programs will want to use other sampling approaches or have larger sample sizes. See the Annex 3: Calculating Sample Sizes for Various Scenarios: Formulas and Examples, for resources that will enable you to select different levels of precision and/or larger samples sizes.

PART II: BASIC SAMPLING TERMINOLOGY[22]

The following table introduces key terms that we use in this introduction and in Modules 2 and 3. We will boldface terms from this table as they are introduced in the text that follows. We recognize that having all these terms listed here without context can be overwhelming, but we encourage you to use this table as a quick reference as you use this manual.

[22] This table is adapted from the KPC Training Materials. You can find the reference for and a link to these materials in Annex 1: Key Resources.

Table 2.3: Key terms and concepts

Term/concept	Definition
Sample:	A group of sampling units (such as individuals or households) selected from the target population.
Random sample:	A method of selecting a sample to ensure that each unit in the population has a known chance of being selected (if it is a simple random sample, each unit has the same probability of being selected).
Random number:	A number that is selected (by chance) from many numbers. Each number has a known chance of being selected.
Cumulative:	Increasing a sum by continuing to add to it. For example, assume that there is a list of three communities. Community A has 40,000 people, Community B has 60,000 people, and Community C has 50,000 people. The cumulative population of Community A and Community B is 100,000 (40,000+60,000). The cumulative population of Community A, Community B, and Community C is 150,000 (40,000+60,000+50,000).
Systematic sampling:	A sampling approach that involves calculating a sampling interval based on the required sample size. A random starting point is chosen and then cases are selected from the sampling frame at a sampling interval. In this approach, the selection of the first sampling unit determines completely the selection of the remaining sampling units in a given frame. Systematic sampling is typically (though not always) used for cluster and LQA sampling methodologies.
Sampling interval:	The total population size (N) divided by the sample size (n). Used as part of systematic sampling to select units from a sampling frame.
Multi-stage sampling:	A process involving more than one step of sampling before reaching the ultimate unit of interest. For example, with cluster sampling, projects first sample clusters from a sampling frame covering the entire target population, then households within clusters and, finally, mothers/caregivers within sample households.
Strata:	Strata are mutually exclusive groups of sampling units constructed from the list of sampling units before sample selection. Examples of these groups are: types of residence, social-economical zones, ethnic groups, administrative units, In a stratified sample, the sampling error depends on the population variance existing within the strata but not between strata. For this reason, it is important to create strata with low internal variability (or high homogeneity). ATTENTION! Another term that you may come across that is related to but different from strata is “survey domain”, which is a subpopulation or sub-geographic area for which separate estimates are to be provided. Survey domains and strata could be the same but they are not always. Stratifying reduces the sampling error by ensuring that important sub-divisions of the population of study are appropriately represented in the sample. But it does not guarantee enough cases for allowing separate estimation for each stratum. In the case of LQAS, the strata are not based on population characteristics but on residence within a supervision area. And just as in LQAS, stratified random sampling does not always allow for a separate estimate at the strata level. A survey domain could consist of one or several lower-level sampling strata. For example, survey domains could be the first-level stratum in a multi-level stratification. In rapid surveys, if the sample within each strata are large enough to provide estimates, strata and survey domains are often the same thing.

Term/concept	Definition
Stratification:	When there are subgroups with important differences we want to make sure that some of these groups are not under-represented in the sample. Stratification reduces this possibility. Stratification is a process by which the survey population is divided into subgroups or strata that are as homogeneous as possible using certain criteria (i.e. ethnic group; urban/rural; mountain/plain). From a statistical perspective the main benefit of stratification is to reduce sampling errors caused by under-representation of certain subgroups. In addition, if enough respondents from each subgroup are included in the sample, we can obtain coverage estimates for subgroups as well as for the entire project area. However, in this case, we need to weight the coverage estimates from all subgroups to get a valid overall estimate. Stratification can be single-level or multi-level. A typical two-level stratification in large surveys is region crossed by urban-rural stratification. A possible rapid survey two-level stratification could be sub-districts crossed by ethnic group belonging (if can be indentified before sampling), or by type of residence.
Sampling frame:	Complete list of every possible sampling unit within the target population from which a sample will be drawn. In multi-stage sampling one would have a different sampling frame for each stage: community, household, and individual respondent, for example.
Enumeration Area (EA):	An EA is a convenient counting unit created for a population census. A typical EA is a small geographic area with clearly delineated boundaries in which typically 100-300 households reside. In rural areas, an EA is usually a village, or a part of large village, or a group of small villages; in urban areas, an EA is often a city block. A list of EAs can serve as a sampling frame in multi-stage sampling. EAs are often defined as part of a national census. If available they can be used to construct the sampling frame. This is especially useful for urban areas An EA frequently serves as a “primary sampling unit in two stage cluster sampling.
Sampling unit:	It is the unit selected through a sampling procedure. Alternatively, this is referred to as the eligible sampling unit. For rapid surveys, a sampling unit is usually the individual or the household, and—see below—frequently corresponds to the ultimate sampling unit from which information will be collected.
Elementary unit:	The ultimate sampling unit to which a survey collects information. In some cases, as in larger surveys, the first stage or second stage sampling unit could be the household, selected from the sampling frame, containing different individuals (i.e. mothers), who will be the final respondents. In this case the individuals are the ‘elementary unit’ within the higher level sampling unit. In Rapid Surveys, the sampling unit and the elementary unit are frequently the same thing. Alternatively, this is referred to as the ultimate sampling unit.
Sample size:	Number of units (individuals, households) selected from the population for inclusion in a study.
Probability proportionate to size (PPS):	A sampling principle that ensures that the sample’s distribution mirrors the population’s distribution. Communities with larger populations have a proportionately greater chance of having clusters or interview sites located in those communities than communities with smaller populations.
Cluster sampling:	A method of sampling population clusters first rather than individuals and then interviewing a certain number of individuals (or all of them) within each cluster to achieve the desired sample size. Cluster sampling is a form of multi-stage sampling aimed at increasing field work efficiency, but may introduce a design effect which usually increases the sampling errors.
Cluster:	A naturally occurring group of individuals from a population of interest. For example, a cluster can be a census EA provided by the population census, with a measure of size equal to the number of households or the population in the EA. Rapid surveys rarely have access to this information.

Term/concept	Definition
Design effect:	Measures the efficiency of the survey design compared to Simple Random Sampling. In cluster sampling, we can calculate the design effect during analysis for each indicator and use the resulting number to adjust the confidence intervals. It can also help adjust sample size in subsequent survey design.
Lot Quality Assurance Sampling (LQAS):	A special form of stratified sampling that allows projects to identify areas with levels of coverage that are at or above expectation versus those that are below expectation
Supervision area:	A subset of the population managed by specific health staff sampled by LQAS methodology to identify staff performance and for project management. ATTENTION! Supervision areas are not clusters—there is no overlap between the two concepts. In fact supervision area is a management concept, used to identify strata for stratified random sampling in LQAS. (LQAS being a specific form of stratified random sampling.)
Precision:	A term that refers to the magnitude of sampling errors or the range of possible values for a given estimate. For example if an estimate of 45% is obtained from the survey, a 10% (absolute error) level of precision means that the true value could be from 35% to 55%. A greater level of precision, for example 5% means that the true value could be from 40% to 50%.
Standard error:	Also known as sampling error which is the square root of the variance of an estimator. It is a statistical measure that indicates the precision of a sample estimate and is used to calculate the confidence limits of that estimate. Typically, the 95% confidence limits of an estimated are calculated as the estimated value minus and plus two times of its standard error.
Confidence interval (limits):	Indicates the range of possible values within which the sample estimate will fall in a certain percentage of the time for all possible samples of identical size and design. Confidence limits are the highest and lowest values within that range and are usually calculated at a level of 95%. That is, there is a 95% chance that the actual rate or proportion being estimated in the study falls within the confidence interval. Confidence intervals indicate the level of precision of our sample estimation (see the definition of precision above).
Bias:	An error that consistently results in an over- or under-estimation of a value of measurement. Bias can be introduced at any step of a survey, from sampling, data collection to data analysis if not properly treated. Use of a random and/or systematic sampling process may help prevent "selection bias." Reducing non-sampling errors (all kinds of errors except sampling errors) to minimum is the best way to prevent bias.

Part III: Sampling Methodologies

Simple Random Sampling

Simple random sampling is the most straight forward sampling design. It is often used as a point of comparison for other sampling methods, especially when determining the **design effect.** Generally, simple random sampling requires a smaller **sample size** than other methods, such as **cluster sampling**.

In simple random sampling each **sampling unit** (e.g., all mothers of children 0–23 months in the project area, all households in the project area) has an equal chance of being included in the survey. What makes this simple random sampling is the methodology for accomplishing this. It involves first making a list of all sampling units in the program area (a **sampling frame**), then assigning each sampling unit a unique number, and finally randomly selecting units until you reach the number that you need for your survey. It is simple to explain and easy to see how the information can be generalized to the population of the program area.

When constructing the sampling frame, we must be sure that all sampling units (i.e. households) in the survey area are included. One can make a sampling frame for SRS by conducting a census of all sampling units (often households) in the entire program area. However, it is better to use a list that has already been generated. For example, if another program is already operating in the area and has conducted a census then this list can be used for the survey. When using a sampling frame that has already been generated, make sure that it is sufficiently up-to-date, so that there is not a lot of difference from current situation. For SRS there is only one sampling stage and therefore only one sampling frame. For **multi-stage sampling** designs, you will create a sampling frame for each stage (more details are explained in sections on cluster sampling.)

There are a variety of ways that a simple **random sample** can be taken from the sampling frame. The names of all sampling units (i.e., households) could be put into a hat and be pulled out randomly, but this would be very time consuming given the population sizes of most programs or projects. There are more practical techniques. For example, one technique is to number the sampling units from 1 to the end of the list. Then a list of **random numbers** could be generated either by consulting a random number table or using random numbers generated by computer programs. The various training materials referenced in Annex 1: Key Resources include examples of random number tables.[23]

Although SRS is simple to explain and often requires a smaller sample size than other sampling methods, there are several factors that make it impractical for rapid surveys:

- A household list may not be available for the entire program area, and it may not be feasible to conduct a census. This is very often the case.
- Performing the random selection of each sampling unit might be overly time-consuming for large populations.
- If analyses are desired for certain sub-groups, such as children 0–5 months or specific ethnic groups, then SRS may not result in large enough samples of these groups. Other sampling methods more efficiently address this situation.
- Often implementing SRS is more costly than other methods.

For these reasons, we will not discuss simple random sampling further.

Stratified Random Sampling

Stratified random sampling regroups similar units into stratum before sample selection, and then designs independent samples for the **strata. Stratification** reduces sampling errors and allows for information to be collected from sub-groups (strata) in order to perform analyses for

[23] Another method, using an Excel spread sheet is described in the Research Methods Knowledge Base. This involves first listing all sampling units in one column. In the column next to it paste the function =RAND() which is Excel's way of putting a random number between 0 and 1 in the cells. Then, sort both columns —the list of names and the random number—by the random numbers. This rearranges the list in random order from the lowest to the highest random number. Then take the first 300 (if a sample size of 300 hundred is desired) in the sorted list. This is efficient if the sampling list is already computerized.

these groups. **LQAS** is a form of stratified sampling where **supervision areas** are strata. (We explain supervision areas in greater detail in the section on LQAS.) Stratified sampling also can be used with two-stage cluster sampling. For example if the survey area contains both rural and urban sub-areas and we want information from both sub-areas, then stratified sampling is useful. Another example is if we have different ethnic groups and want to be sure that we have representatives from each group in the sample. Then, in this situation, stratified sampling is appropriate.

Imagine that in the earlier example, of the survey that is conducted in a mountainous area, there are three different languages spoken. The most common language is Spanish (70%); however, 20% speak Aymara and 10% speak Quechua. Suppose we are conducting interventions to increase the use of skilled birth attendance at deliveries in the program areas. We have reason to suspect that the level of skilled attendance at birth is different for each of the language groups. Suppose that we determined that the minimum sample size needed from each sub-group is 100 to look at sub-group coverage levels. If we take a simple random sample, we may not get enough interviews from the Aymara and Quechua groups. We note in passing that the choice of sample size is related to our desired **precision**—that is what is meant by the concept of "needed from each sub-group." We want a certain level of precision for estimates of proportion of births attended by skilled attendants in each language group.

While simple random sampling may not provide us with enough respondents from each language, group stratified random sampling ensures that we have a large enough sample in each sub-group. How does this work? As with SRS, you need a list of sampling units (a sampling frame). In addition, for every sampling unit (i.e., household or individual), you need information about the characteristic that you are using to divide the population (e.g., language spoken). This way, the population can be divided into separate groups or strata. Strata are mutually exclusive groups of sampling units from the list of sampling units. In this example, we would have to restrict inclusion in language groups, for example by specifying primary language spoken, so that the population can be divided into mutually exclusive groups. We assemble lists of sampling units for each sub-group, and then we select respondents using simple random sampling in each sub-group. For this example, we can sample 100 respondents from each of the three language groups. We can determine coverage levels for each language group and combine them, using a weighted average, to determine coverage for the entire program area. We discuss weighting in detail in Module 3, but it simply means adjusting the total estimate by accounting for the relative size of each language group in the total population.

Although this method solves the problem of having a large enough sample from each sub-group, it retains some of the problems of SRS:

- A household list may not be available for the entire program area, and it may not be feasible to conduct a census. In addition, if a list does exist, it may not contain information about the characteristic needed for stratification (e.g., primary language).
- Performing the random selection of each sampling unit might be overly time-consuming for large populations, and SRS has to be performed separately for each sub-group.

The following tables illustrate the difference between SRS and Stratified Random Sampling.

Table 2.4: Sampling using SRS and stratified random sampling

Record no.	Name	Language	Six respondents chosen using SRS
1	Aparicio	Spanish	
2	Apaza	Aymara	X
3	Gomez	Spanish	X
4	Hilari	Aymara	
5	Mallo	Spanish	X
6	Mamani	Quechua	
7	Mejia	Spanish	
8	Monasterios	Spanish	
9	Morales	Aymara	X
10	Paredes	Spanish	
11	Paxi	Quechua	
12	Quispe	Aymara	
13	Reyes	Spanish	
14	Riojas	Spanish	
15	Rodriguez	Spanish	
16	Romero	Spanish	
17	Torrico	Spanish	X
18	Santander	Spanish	X
19	Villa	Spanish	
20	Yanez	Spanish	

Stratified Random Sampling

Record no.	Name	Language	Six respondents (two/stratum) chosen using stratified random sampling
6	Mamani	Quechua	X
11	Paxi	Quechua	X
2	Apaza	Aymara	
4	Hilari	Aymara	X
9	Morales	Aymara	
12	Quispe	Aymara	X
1	Aparicio	Spanish	
3	Gomez	Spanish	
5	Mallo	Spanish	
7	Mejia	Spanish	X
8	Monasterios	Spanish	
10	Paredes	Spanish	
13	Reyes	Spanish	
14	Riojas	Spanish	
15	Rodriguez	Spanish	
16	Romero	Spanish	
17	Torrico	Spanish	
18	Santander	Spanish	
19	Villa	Spanish	
20	Yanez	Spanish	X

Results

Group	No. chosen SRS	No. chosen stratified
Quechua	0	2
Aymara	2	2
Spanish	4	2

SRS resulted in no Quechua respondents being selected, so that it would not be possible to make estimates about coverage for this group. However, with stratified sampling, each sub-group is represented in the survey by design so that we can develop estimates (with a given level of precision) for each sub-group.

Systematic Sampling

Systematic sampling is a method for choosing a sample from a list that is easier than SRS. Understanding this method provides the basis for understanding **PPS**[24] sampling, which is used in cluster sampling and LQAS.

Systematic sampling involves selecting a sample by choosing every k^{th} sampling unit in the study population. The value for k is the **sampling interval.** We determine the sampling interval by using both the sample size needed and the total number of sampling units in the population. The advantage of systematic sampling is that one needs only one random number. The first step to accomplishing systematic sampling is to determine the following:

- Number of sampling units in the population from *1 to N* (*N* is the total number of sampling units)
- Sample size or *n* needed
- Interval size or *k,* defined as k=N/n
- A random number between 1 and *k* (you might call this the "random start")

For the second step, you need a list of the sampling units. If this list is in random order, systematic sampling approximates SRS. If the sampling units are grouped by a characteristic (strata) then systematic sampling is basically the same as stratified sampling.

Third, you choose the first sampling unit to be included in the study. This corresponds to the random number—the random start—between 1 and *k*. Fourth, you choose the rest of the sample by counting down the list and choosing every k^{th} sampling unit. This results in choosing the number of sampling units equal to the desired sample size (*n*).

[24] PPS sampling is a sampling principle that ensures that the sample's distribution mirrors the population's distribution. Communities with larger populations have a proportionately greater chance of having clusters or interview sites located in those communities than communities with smaller populations. More details will be provided in Module 2.

The following illustrates this process:

List of sampling units

1	26	51	76
2	27	52	77
3	28	53	78
4	29	54	79
5	30	55	80
6	31	56	81
7	32	57	82
8	33	58	83
9	34	59	84
10	35	60	85
11	36	61	86
12	37	62	87
13	38	63	88
14	39	64	89
15	40	65	90
16	41	66	91
17	42	67	92
18	43	68	93
19	44	69	94
20	45	70	95
21	46	71	96
22	47	72	97
23	48	73	98
24	49	74	99
25	50	75	100

- Suppose that you have a total population of 100 (N) and have determined that you need a sample size (n) of 20.
- The interval size (k) is N/n or 100/20=5.
- You choose a random number between 1 and 5. Suppose that this number is 4. This is your random start.
- You start the selection process by first choosing the 4th sampling unit on the list.
- You choose the second sampling unit by adding the interval number 5 (k) to 4 to get the 9th sampling unit from the list.
- You continue this process until you reach the sample size of 20, which will also be toward the end of the list. In this example the 99th sampling unit is the last chosen.

We use systematic sampling routinely in both cluster and LQAS usually to select communities (more correctly, clusters or interview sites) in the first stage of our sampling process. In such cases, we have a list of communities and their populations in a program or supervision area and use systematic sampling to select clusters or interview sites.

Cluster Sampling

This is a brief introduction to cluster sampling because we develop it in more detail in Module 2. Cluster sampling is a good choice for rapid surveys. This method reduces both time and cost because there are fewer sites to visit than in simple random sampling. In cluster sampling, we do not directly select sampling units (e.g., mothers of children 0–23 months). Instead we first randomly select groups (clusters) of sampling units; then randomly select sampling units from each of the clusters. A cluster is a naturally occurring group of individuals (such as a village, ward, or city bock) likely to include the population groups your project is interested in studying (i.e. children 0-23 months and their caregivers). For rapid surveys, it is common to select 30 clusters. Field costs and time spent collecting data are often greatly reduced by only having to visit 30 sites instead of more than 100 sites that might be required for SRS sampling.

A set number of sampling units are usually selected from each cluster. This number varies, depending on what is being studied. Expanded Programs of Immunization have traditionally used a 30x7 cluster sampling design. In this design, one selects seven sampling units from each of the 30 clusters. Health programs that measure a variety of interventions often use 30x10 cluster designs (10 sampling units from each of 30 clusters for a total of 300 respondents—see text box), while surveys of nutrition programs often use a 30x30 design. There are cluster designs where all sampling units in each selected cluster are interviewed, but this is not usually the case for rapid surveys of public health interventions (although it is the case for some national surveys such as the DHS). In the rest of this handbook, we will use the 30x10 cluster sampling method for our examples.

A possible source of confusion about sampling units

This manual emphasizes methods where the number of respondents to the survey questionnaire is per-determined in the design. In the case of 30x10, there will be 10 respondents in each of the 30 clusters. Respondents, found in household, are the ultimate sampling unit.

Readers will also come into other resources where the main sampling unit is the household, where surveyors will interview all eligible respondents. In that case the number of households is pre-determined. This is the approach followed by the DHS survey. The pros and cons of each method and why we will stick to the former is discussed in a special section on caveats and potential biases.

Cluster sampling is a multi-stage sampling method, where one stage is the selection of clusters. For this stage your sampling frame is a list of clusters with approximate population for each cluster. For example, the population for each cluster could be total population, number of households, or number of children 0-23 months in the survey area. Suppose that clusters are villages and you have information on the number of children 0–23 months for each village. You prepare a sampling frame composed of a list of all villages and the number of children 0–23 months for each village in the survey area. Please make sure that all households can be assigned to a village. Do not leave anyone out of the sampling frame. For urban areas a sampling frame could be made from a list of city blocks or neighborhoods and their populations, again making sure that no one is left out.

Cluster sampling methods have to take into consideration and actually calculate the design effect that results from this sampling method. The design effect is the result of the tendency of sampling units within one cluster to be more like each other than like sampling units from other clusters. Cluster sampling can be compared to SRS, which has a design effect of 1. Because cluster sampling has a higher design effect than SRS, it requires a larger sample size to obtain the same level of precision. There are formulas for calculating design effects based on the relative homogeneity within clusters as compared to between clusters. You can only use these formulas after you have collected the data; however, rapid health surveys often use a rule of thumb and estimate the design effect to be 2 for purposes of sample size determination. The result of this example is that if a sample size of 150 is required for SRS, then cluster sampling requires a sample size of 300, which can be obtained through a 30x10 cluster design. You can obtain larger or smaller sample sizes by either increasing the number of sampling units chosen in each cluster or increasing the number of clusters.

The advantage of cluster sampling is that it typically helps you reduce the costs and time required to conduct a survey. However, there are disadvantages. The main disadvantage is that cluster sampling only provides information for the entire survey area. It is not possible to determine which supervision areas are more successful than others without needing a very large sample size. The following is an example.

> *Suppose that the entire project area is a district. The district is divided into 5 sub-districts each with a medical doctor in charge. Each of these 5 sub-district heads wants information on how his or her district is performing. We cannot provide any information at the sub-district (supervision area) level unless we take a large enough sample in each sub-district and this is likely to be prohibitive in terms of cost and time. Moreover, it is possible for the overall coverage of an indicator, such as skilled attendance at birth, to be high for the entire project area at the same time that one or two of the supervision areas are performing poorly with regard to that indictor (i.e., with almost no women receiving skilled birth attendance). These unsuccessful supervision areas are masked by the rest of the supervision areas that are successful (i.e., where skilled birth attendance is common) unless we take a large enough sample in each sub-district. Again, this is costly to do. If it is important to have some information about supervision areas within the project area, then LQAS is recommended.*

Lot Quality Assurance Sampling

LQAS is form of **stratified sampling** that gives us some information about what is happening in the sub-areas and also gives us coverage estimates for the entire project area while maintaining a relatively small sample size for the entire program area. For example, if you are a program manager for a health district that is divided into five administrative areas, you may want to know which of the administration areas are on target for Tetanus Toxoid (TT) coverage of pregnant women. Suppose you have a target of 80% coverage for TT. If you perform a survey using the cluster method, you may find that coverage for the entire district is 80%; however, you may suspect that one of the administrative areas is not actually meeting the target of 80%. The survey using the cluster method does not provide any information about these administrative areas, so you do not actually know how they are performing. However, LQAS solves this problem by providing yes/no information about whether a sub-area (often referred to as a supervision area [SA]) is reaching the target. The following is a brief overview of the basic steps for LQAS:

- Divide the program area into SAs that have meaning for program management. For this example, there are five SAs.
- In each SA, using PPS, select a set number of interview sites, usually 19 (we discuss the rationale for this number in Module 3.)
- Conduct one interview in each site:
- Surveys that collect information from more than one sampling group will use parallel sampling at each interview site. We discuss this issue in detail in Module 3.
- Analyze information for each SA using a decision table that allows you to determine whether the target was met based on the number of correct responses in each SA.
- Combine information from each SA using weighted averages to determine the coverage estimate for the entire program area.

We will not expand on this information at this point, since we provide details in Module 3. However, please note that LQAS is a stratified sampling method, with each SA as a stratum. We want to be able to say something about each SA—each stratum—and so, by design, we select a sample of 19 in each SA. To develop an estimate for the entire program area, we then weight the results by using the relative size of each SA to the total population as the weight. We will provide examples of how to do this in Module 3.

LQAS is a good choice for managers who need to make decisions about sub-areas. You can analyze information using spreadsheets instead of complex computer programs. Also, similar to cluster sampling, LQAS can provide coverage information for the entire project area.

However, there are disadvantages to this methodology as compared with cluster sampling. The main difficulty is that you have to visit more interview sites than in cluster sampling, which may increase the cost and time needed for data collection. In the above example there are five SAs, and 19 interviews are performed in each SA. This means that you will have to visit 95 (5*19) sites instead of 30 interview sites that is common for cluster sampling. Another disadvantage is that you have to use parallel sampling if the survey collects information from multiple sampling groups (e.g., mothers of children 0–5 months, mothers of children 0–23 months, and adolescent males 15–24 years). In contrast when two-stage cluster sampling is used, usually the sample size is large enough to be able to analyze information from multiple sampling groups, although there are some instances where parallel sampling is used. The difference is that for LQAS parallel sampling is always needed to collect information from multiple sampling groups. If you use parallel sampling, you have to remember to use weighted averages to calculate coverage for the entire program area. More details on parallel sampling for LQAS and how to calculate weighted averages are in Module 3.

Despite these disadvantages, managers often find that it is worthwhile to use LQAS in order to be able to tailor program efforts to the needs of sub-areas, thus making LQAS a good choice for them.

PART IV: ADDITIONAL GENERAL SAMPLING TOPICS

Choosing between Cluster and LQAS

At this point, you probably do not have enough information to determine which sampling methods are most appropriate for your program and your rapid survey. We would encourage you to carefully read Modules 2 and 3 before making a determination. However, we have already suggested some advantages and disadvantages of LQAS vis-à-vis cluster sampling and would invite you to consider them as you read on.

To put it simply, if you have several distinct SAs for which different managers or project staff need their own local information to assess (yes or no) whether they are meeting program targets, then you might consider using LQAS. If, on the other hand, you feel that having coverage estimates for the entire program area is sufficient, you might consider using a 30x10-cluster sample.

Modifications for Urban Areas

Whether you are implementing a rapid population based survey in a rural area or an urban area, you still follow the same steps of preparing for the survey, training staff, carrying out the field work, and using the results. You can use both two-stage cluster sampling and LQAS in urban areas. As you plan the survey, you still determine why you are doing the survey, whom you will survey, what you will ask them, how you will find respondents, who will carry out the survey, and how much it will cost (based on logistics). However, you will find differences in the details regarding how to find respondents and how much it will cost.

The first adjustment you make is how you create the sampling frame. In rural areas you often use villages as the sampling unit. However, urban areas are not usually comprised of villages. Instead you can use **enumeration areas** from a national census, neighborhoods or city blocks. Choose a list of sampling units that is available and logical to use for the urban area where you will be conducting the survey. Unless these sampling units have equal population sizes, you also need an estimate of the population in each sampling unit. At that point, PPS sampling is the same for both urban and rural areas.

The second adjustment is if there are multifamily dwellings (i.e. apartment buildings or compounds where many families live) in the survey area. In this case, you must carefully plan for choosing households. The following are two procedures for household selection:

Table 2.5: Choosing households in multifamily dwellings

More than One Story (more than one floor)	Single Story (one floor)
• Randomly select one floor (Using a random number table?)	
• Number all the households on that floor	• Number all the households in the dwelling
• Select a random number between 1 and the number of households on the floor	• Select a random number between 1 and the number of households in the dwelling
• Find the household on the numbered list whose number matches the random number you selected	• Find the household on the numbered list whose number matches the random number you selected
• Go to the household	• Go to the household

The third adjustment is that you must understand the work habits of the respondents and determine the best time to visit households for interviews.[25] Are respondents away from their homes during the day (i.e. because the sell goods in the local market or they work in factories during the day)? If they are not normally home, then interviews should be scheduled for the evening when they are home.

Fourth, you must make adjustments to survey costs and logistical arrangements. Often because distances are not as great as in rural areas, logistics costs are lower for urban areas. Interviewers and supervisors may be able to use public transportation instead of a vehicle specifically dedicated to the survey. However, it is important to make sure that public transportation is

[25] Although this is also true for rural areas, it is an especially necessary step for urban surveys because it is very common for respondents in urban areas to work away from home during the day.

available for all areas where interviews will be carried out and for the time of day when interviews will be conducted. It may be most practical to use a combination of public transportation and one dedicated vehicle. Special consideration must be made for safety of interviewers, especially if they are interviewing at night or in high crime areas.

Fifth, you must train supervisors and interviewers in household selection and logistic plans. Supervisors and interviewers must know the limits of sampling units (enumeration areas, city blocks or neighborhoods) in order to be sure that they are properly selecting households.

Parallel Sampling

In the first section of the Section II Introduction, we discussed the need to take stock of what information you need and illustrated how to reduce the number of groups sampled, even if our program focuses on several groups. Even if we can reduce the number of different groups to be sampled, it is still very common that we will have more than one group—more than one population of interest. This could be as simple as needing to distinguish mothers of children 0–5 months (to assess exclusive breastfeeding) from mothers of children 6–23 months (to assess appropriate feeding). It could be as complex as needing to interview mothers about reproductive issues and young men about sexual practices.

In any case, we often find ourselves interested in different populations. To ensure that we find enough respondents for each indicator of interest (given that different indicators might require different kinds of respondents), we can adopt an approach (after we have selected **clusters** or interview sites) known as parallel sampling. We provide specific guidelines for how to conduct a parallel sampling for LQAS in Module 3 (because it is a particular challenge for LQAS due to its relatively small total sample). However, here we lay out general steps for parallel sampling, which are common to a variety of sampling methods including LQAS and two-stage cluster sampling.

Steps for parallel sampling (assuming you have already selected the interview site or community):

1. Determine the number of distinct sample groups.[26]
2. Develop separate questionnaires for each sample group.
3. Select the first household (using procedures we discuss in Module 2).
4. Administer all appropriate questionnaires in that household (if there are two children of different ages implying two separate questionnaires, for example, it is appropriate administer both questionnaires to the same caregiver—see example below).
5. Once eligible respondents have been interviewed in the first household, select the second household in which an interview can occur. We develop several methods for selecting the second and subsequent households in Module 2 on cluster sampling. Administer any remaining questionnaires as appropriate.
6. Continue, as needed, to the next household using the agreed upon protocol for selecting the next household and administer any remaining questionnaires as appropriate.
7. Continue in this way until all questionnaires have been filled out.

[26] See above section: Take stock of what information you need.

AN EXAMPLE OF PARALLEL SAMPLING

Your project is interested in the following population groups:

- Non-pregnant women 15–49
- Mothers of children 0–11 months
- Mothers of children 12–23 months
- Men 15–49

Let us assume you are conducting a 30x10 cluster sample, and that the Core Team has decided to use parallel sampling. In each cluster, it wants to conduct 10 interviews for each of the above groups, and it has designed a separate questionnaire for each. The two questionnaires targeting men and women of reproductive age only contain questions on knowledge of child spacing methods, HIV/AIDS, and sexually transmitted diseases. One of your interviewers visits a household where there are two sisters and their sister-in-law living in the same household. Each woman has a husband who is between 15 and 49 and also lives in the household. One of the women is 19 and has no children. Her sister is 21 and has a child who is 18 months. Their sister-in-law has a baby who is 3 weeks of age. None of the women are currently pregnant. What should the interviewer do? The interviewer can administer three questionnaires in that household:

- One questionnaire for women 15–49
- One maternal questionnaire (either the one for mothers of children 0–11 months or the one for mothers of children 12–23 months)
- One questionnaire for men 15–49

Any of the women can answer the questionnaire designed for women 15–49. However, the interviewer will need to randomly select one of the two mothers in the household and then administer the correct questionnaire based on the age of her child. Can the same woman answer questions from both the 15–49 questionnaire and the maternal questionnaire? Yes, if the two questionnaires do not contain the same questions, or the project will not be aggregating the data from non-pregnant women and mothers of young children. In those instances, it would be acceptable to administer the two questionnaires to the same woman. In terms of the men's questionnaire, the interviewer can randomly select one of the husbands for the interview.

Once the interviews are completed, the interviewer can go to the next household (using the agreed upon method for selecting a second household). The interviewer and his/her team members should continue visiting households in that sample area until they have 10 interviews in each of the four groups. Once the survey team has completed the required number of interviews for a particular group, it can focus on getting the required number of interviews in the remaining groups. For example, assume that the survey team has completed 10 interviews with women 15–49, 10 interviews with mothers of children 0–11 months, 7 interviews with mothers of children 12–23 months, and 4 interviews with men of reproductive age. For the remaining households in that area, the team should only sample mothers of children 12–23 months and men 15–49. Once the team has completed 10 interviews in each of those groups, it should move on to the next sample area.

MODULE 2: TWO-STAGE CLUSTER SAMPLING (30X10 DESIGN)

INTRODUCTION

In the previous section on sampling principles and terminology, we introduced the concept of cluster sampling and briefly compared it with simple random and LQAS methodologies. In this section, we walk through the process for using a two-stage 30x10 cluster sampling approach for your rapid survey.[27] Recall that 30x10 means that you will choose 30 clusters within your program area and conduct 10 interviews in each cluster (recalling that if you have multiple questionnaires, as in the case of parallel sampling, that you will conduct 10 interviews for each questionnaire). Let us examine the procedures for selecting a sample first before looking at more technical information related to calculating design effects and using alternative sample sizes.

PRINCIPLES OF TWO-STAGE CLUSTER SAMPLING

In the Section II Introduction, we presented cluster sampling as an alternative to simple random sampling that is more efficient in terms of time and cost. Recall that a cluster is a naturally occurring group of individuals (such as a village, ward, or city block) that includes the population group your project is interested in studying (e.g., children under 24 months and their caregivers). Cluster sampling has become widely used for rapid surveys for two primary reasons:

1. It requires (in the first instance) a sampling frame that is merely a list of population centers (such as towns, villages, or communities) and their estimated population sizes. Such a sampling frame is usually easy to obtain even in resource-poor environments.
2. Interviewing a number of people who live in the same cluster reduces time and travel costs between interviews.

Because of the first point above, cluster sampling requires random selection of sampling units at two (or more) stages, thus the moniker: two-stage cluster sampling. [28] In the approach described here we have two stages:

- Stage 1: Sample communities from the population to get sample areas (clusters)
- Stage 2: Sample households within sample areas (in which we select final sampling units (e.g., primary caregivers of children under 24 months of age)

This means that we take two random samples: a random sample of communities to select clusters and then a random sample of households within the cluster to find respondents (sampling units). Some resources would say that there are actually three stages because you might find a

[27] 30x10 cluster sampling is a cost-effective widely understood sampling approach, but both the number of respondents per cluster and the number of clusters can be increased. Please see additional statistical references, in Annex 1: Key Resources.

[28] "Two-stage" cluster sampling refers to the two main stages described here. There are cases, some described in this manual, where further sampling stages take place, for example within a household. In those cases, "two-stage" becomes a misnomer and should be replaced by "multi-stage" sampling. We find that it is a convenient short-cut for the type of multi-stage sampling described here (aka 30x10) and will continue using two-stage sampling in the rest of the manual.

household in Stage 2 in which there are two or more sampling units and, in such a case, you need to randomly select one of them. While this may be the case (and we will describe what to do in such a case), we will still refer to this approach as "two-stage."

Clusters are usually selected using systematic sampling (which we introduced in the previous section). To use systematic sampling, a project needs 1) a sampling frame of all communities in the program area and 2) a sampling interval. The sampling frame should include every community in the project area, its population size, and its **cumulative** population. Cluster sampling, like simple random sampling, is a probability sampling design. In other words, the selection of sampling units is based on chance. Therefore, it is important to list every community in the program area, regardless of its size. Clusters are selected with PPS. This means that larger communities have a greater chance of having clusters than smaller communities. Why sample with PPS? You do this if you want the distribution of cases (mothers/caregivers and young children) in the sample to mirror the distribution of cases in the general population.

With these concepts and principles in mind, we now proceed to a step-by-step example of how to select a sample using this approach. We start with the process of selecting clusters before turning to the approach to selecting households (and respondents) within selected clusters. Thus we start with the procedure for taking a systematic sample—using a probability of selection proportionate to the size of the communities—to define the clusters we will visit to interview respondents. We list the steps assuming a 30x10 cluster sample design for reasons we already discussed.

STEPS FOR SELECTING CLUSTERS WITH PROBABILITY PROPORTIONATE TO SIZE

Here are the steps, which are described in a detailed example in the pages that follow:

1. Create a list (ordered randomly) of communities (villages/towns/wards), the population for each, and the total population of all communities. Refer to existing population data to get the size (number of residents or households) of each community in the program area.
2. Calculate the cumulative population of each community by summing the total population of the community with the combined total population of all the preceding communities on the list. NOTE: The cumulative population of the last community listed in the sampling frame should equal the total population of the entire program area. If this is not the case, check your calculations.
3. Determine the sampling interval by dividing the total population of the entire program area by the total number of clusters you desire to select (30, in our case).
4. Choose a random number to identify the starting point on the list to begin selecting clusters. The random number must be less than or equal to the sampling interval.
5. Beginning with the random number, use the sampling interval to identify communities for the 30 clusters.

Suppose we have a program area in which there are 50 communities where we have been carrying out the program. We will use the same table below to show each step in the process.

- **Step 1:** Table M2.1 has a listing of all communities with their populations and the total program-wide population of 301,170.

- **Step 2:** Table M2.2 calculates the cumulative population by adding each community's population to the total of all communities before it. Verify the numbers. Notice that the cumulative population of the last community—Varok—is equal to the total population.

Table M2.1: Step 1: List of communities and their populations

Name	Population	Cumulative Population	Clusters	Number of Clusters
Utaral	12888			
Bolama	3489			
Talum	6826			
Wara-Yali	4339			
Galey	2203			
Tarum	4341			
Hamtato	1544			
Nayjaff	885			
Nuviya	2962			
Cattical	4234			
Paralal	1520			
Egala-Kuru	3767			
Uwanarpol	3053			
Hilandia	60000			
Puratna	2297			
Kagaini	1355			
Hamali-Ura	833			
Kameni	4118			
Kiroya	2802			
Yanwela	3285			
Bagvi	4396			
Atota	3188			
Kogouva	1179			
Ahekpa	612			
Yandot	3193			
Nozop	17808			
Mapasko	3914			
Lothoah	15006			
Voattigan	9584			
Pliotok	13225			
Dopoltan	2643			
Coccopa	26000			
Famezgi	3963			
Jigpelay	2115			
Mewoah	507			
Odigala	3516			
Sanbati	14402			
Andidwa	2575			
Ore-Mikam	3105			
Dunu-Mikam	4176			
Kedi-Sina	1919			
Panabalok	3261			
Rokini	4270			
Talosso	3301			
Djaragna	3250			
Bibachi	4670			
Bilam	757			
Sisse	12037			
Anda-Dali	2155			
Varok	3702			
Total	**301170**			

Table M2.2: Step 2: Cumulative population of all communities

Name	Population	Cumulative Population	Clusters	Number of Clusters
Utaral	12888	12888		
Bolama	3489	16377		
Talum	6826	23203		
Wara-Yali	4339	27542		
Galey	2203	29745		
Tarum	4341	34086		
Hamtato	1544	35630		
Nayjaff	885	36515		
Nuviya	2962	39477		
Cattical	4234	43711		
Paralal	1520	45231		
Egala-Kuru	3767	48998		
Uwanarpol	3053	52051		
Hilandia	60000	112051		
Puratna	2297	114348		
Kagaini	1355	115703		
Hamali-Ura	833	116536		
Kameni	4118	120654		
Kiroya	2802	123456		
Yanwela	3285	126741		
Bagvi	4396	131137		
Atota	3188	134325		
Kogouva	1179	135504		
Ahekpa	612	136116		
Yandot	3193	139309		
Nozop	17808	157117		
Mapasko	3914	161031		
Lothoah	15006	176037		
Voattigan	9584	185621		
Pliotok	13225	198846		
Dopoltan	2643	201489		
Coccopa	26000	227489		
Famezgi	3963	231452		
Jigpelay	2115	233567		
Mewoah	507	234074		
Odigala	3516	237590		
Sanbati	14402	251992		
Andidwa	2575	254567		
Ore-Mikam	3105	257672		
Dunu-Mikam	4176	261848		
Kedi-Sina	1919	263767		
Panabalok	3261	267028		
Rokini	4270	271298		
Talosso	3301	274599		
Djaragna	3250	277849		
Bibachi	4670	282519		
Bilam	757	283276		
Sisse	12037	295313		
Anda-Dali	2155	297468		
Varok	3702	**301170**		
Total	**301170**			

- **Step 3:** We can now calculate the sampling interval by dividing the total population by the number of clusters we want in the sample (30 in this case):

301,170/30=10,039

Here the sampling interval is a whole number. When it is not, you may round it to the nearest whole number (rounding .5 up to the nearest whole number). If it had been 10,039.27, you would have used 10,039, but if it had been 10,039.5, you would have used 10,040.

- **Step 4:** At this point we can select a random number to identify the starting place on the community list. Recall this number should be between 1 and 10,039, our sampling interval. There are many ways to select a random number. One way is use a random number table (see Table M2.3 for a portion of such a table to illustrate its use).

In order to use the table, in our case, we need a number with five digits because we need a number between 00001 and 10039. The columns on this table each have five digits, so we can use any row/column combination as long as it is between 1 and 10039. You can, without looking at the table, say "Row 3, Column 4" and then go there. This is highlighted, and you can see that the number highlighted—13045—does not fall into the range 00001 to 10039 (it lies just above it). So you can try again. Suppose you choose "Row 8, Column 10. This is highlighted in blue, and we see in this case the number does fall into the desired interval: 9679. This number becomes our random starting place.

Nothing random about random sampling

As this section shows, following a randomization process is far from the colloquial sense of "random," but requires careful consideration and step-by-step respect for important principles.

It is tempting to find shortcuts, which seem to make sense in the heat of the moment. But you will be trying to describe an **entire** population by talking to essentially **just a few** (even 300 is really a few). This is a bold and risky proposition deserving that we maximize care given to making findings from our sample valid for the entire population. This demands that we watch carefully for both sampling and non-sampling errors.

This part is certainly **not** random.

Table M2.3: Random number table

Column / Row	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	87172	43062	39719	10020	32722	86545	86985	04962	54546	23138	62135	55870	97083	67875
2	28900	50851	30543	89185	16747	95104	49852	26467	58869	79053	06894	23975	34902	23587
3	86248	71156	55044	**13045**	33161	95604	57876	23367	10768	78193	60477	70307	06498	48793
4	10531	51391	41884	69759	32741	70072	01902	96656	90584	59263	49995	27235	40055	20917
5	02481	90230	81978	39127	93335	74259	25856	52838	49847	69042	85964	78159	40374	49658
6	23988	13019	78830	17069	58267	69796	94329	34050	25622	55349	10403	93790	77631	74261
7	37137	47689	82466	24243	10756	54009	44053	74870	28352	66389	38729	80349	50509	56465
8	38230	82039	34158	90149	82948	60686	27962	39306	53826	**09679**	76144	38812	76939	03119
9	98745	08288	19108	84791	58470	59415	45456	44839	86274	25091	42809	56707	47169	95273

- **Step 5** shows us how to select the first cluster using the random start and then to select all subsequent clusters by adding the sampling interval to that number and to each subsequent total. The numbers obtained this way point us to the community whose cumulative population contains that number. Here are the 30 numbers we will select using the random start and sampling interval to make calculations:

Table M2.4: Step 5 (Part 1): Identifying locations using random start and sampling interval

Cluster	Calculation	Interview Location
1	Random Number = Location Number 1	**9679**
2	Random Number + Sampling Interval = Location Number 2	9679+10039=**19718**
3	Interview Location Number 2 + Sampling Interval	19718+10039=**29757**
4	Interview Location Number 3 + Sampling Interval	29757+10039=**39796**
5	Interview Location Number 4 + Sampling Interval	39796+10039=**49835**
6	Interview Location Number 5 + Sampling Interval	49835+10039=**59874**
7	Interview Location Number 6 + Sampling Interval	59874+10039=**69913**
8	Interview Location Number 7 + Sampling Interval	69913+10039=**79952**
9	Interview Location Number 8 + Sampling Interval	79952+10039=**89991**
10	Interview Location Number 9 + Sampling Interval	89991+10039=**100030**
11	Interview Location Number 10 + Sampling Interval	100030+10039=**110069**
12	Interview Location Number 11 + Sampling Interval	110069+10039=**120108**
13	Interview Location Number 12 + Sampling Interval	120108+10039=**130147**
14	Interview Location Number 13 + Sampling Interval	130147+10039=**140186**
15	Interview Location Number 14 + Sampling Interval	140186+10039=**150225**
16	Interview Location Number 15 + Sampling Interval	150225+10039=**160264**
17	Interview Location Number 16 + Sampling Interval	160264+10039=**170303**
18	Interview Location Number 17 + Sampling Interval	170303+10039=**180342**
19	Interview Location Number 18 + Sampling Interval	180342+10039=**190381**
20	Interview Location Number 19 + Sampling Interval	190381+10039=**200420**
21	Interview Location Number 20 + Sampling Interval	200420+10039=**210459**
22	Interview Location Number 21 + Sampling Interval	210459+10039=**220498**
23	Interview Location Number 22 + Sampling Interval	220498+10039=**230537**
24	Interview Location Number 23 + Sampling Interval	230537+10039=**240576**
25	Interview Location Number 24 + Sampling Interval	240576+10039=**250615**
26	Interview Location Number 25 + Sampling Interval	250615+10039=**260654**
27	Interview Location Number 26 + Sampling Interval	260654+10039=**270693**
28	Interview Location Number 27 + Sampling Interval	270693+10039=**280732**
29	Interview Location Number 28 + Sampling Interval	280732+10039=**290771**
30	Interview Location Number 29 + Sampling Interval	290771+10039=**300810**

Let us now plug these locations into Table M2.5 on the next page to identify the communities in which we will interview respondents for each cluster.

Table M2.5: Step 5 (Part 2): Selecting clusters based on selected locations

Name	Population	Cumulative Population	Clusters	Number of Clusters
Utaral	12888	12888	9679	1
Bolama	3489	16377		
Talum	6826	23203	19718	2
Wara-Yali	4339	27542		
Galey	2203	29745		
Tarum	4341	34086	29757	3
Hamtato	1544	35630		
Nayjaff	885	36515		
Nuviya	2962	39477		
Cattical	4234	43711	39796	4
Paralal	1520	45231		
Egala-Kuru	3767	48998		
Uwanarpol	3053	52051	49835	5
Hilandia	60000	112051	59874, 69913, 79952, 89991, 100030, 110069	6, 7, 8, 9, 10, 11
Puratna	2297	114348		
Kagaini	1355	115703		
Hamali-Ura	833	116536		
Kameni	4118	120654	120108	12
Kiroya	2802	123456		
Yanwela	3285	126741		
Bagvi	4396	131137	130147	13
Atota	3188	134325		
Kogouva	1179	135504		
Ahekpa	612	136116		
Yandot	3193	139309		
Nozop	17808	157117	140186, 150225	14, 15
Mapasko	3914	161031	160264	16
Lothoah	15006	176037	170303	17
Voattigan	9584	185621	180342	18
Pliotok	13225	198846	190381	19
Dopoltan	2643	201489	200420	20
Coccopa	26000	227489	210459, 220498	21, 22
Famezgi	3963	231452	230537	23
Jigpelay	2115	233567		
Mewoah	507	234074		
Odigala	3516	237590		
Sanbati	14402	251992	240576, 250615	24, 25
Andidwa	2575	254567		
Ore-Mikam	3105	257672		
Dunu-Mikam	4176	261848	260654	26
Kedi-Sina	1919	263767		
Panabalok	3261	267028		
Rokini	4270	271298	270693	27
Talosso	3301	274599		
Djaragna	3250	277849		
Bibachi	4670	282519	280732	28
Bilam	757	283276		
Sisse	12037	295313	290771	29
Anda-Dali	2155	297468		
Varok	3702	301170	300810	30
Total	301170			

In Table M2.5, we can see that larger communities such as Hilandia actually contain multiple clusters (as more than one cluster will be selected, more respondents will be interviewed). This should not be surprising, given what we said earlier about the use of a selection method with probability of selection proportionate to the size of each community's population. What this implies in Hilandia's case, for example, is that we will need to select—randomly—six clusters within that town/city. This means that we will interview a total of 60 respondents in Hilandia (6 clusters x 10 per cluster). Presumably the city of Hilandia has well-defined sub-units (e.g., "quartiers," wards, or neighborhoods) that you would use to define possible clusters of roughly equal size. You will need to insert a further sampling "stage" in communities with more than one cluster in order to randomly select six clusters (assuming there are more than six neighborhoods/wards/quartiers in Hilandia). You should do this before teams go to the community, and the sampling frame you use in Hilandia should have clear boundaries for each cluster and a random method for selecting six clusters. The simplest way to select clusters in a large community is to list all potential sub-units (neighborhoods/wards) and randomly select the number of clusters needed, choosing random numbers from 1 to the highest numbered sub-unit listed. This will work if the clusters have approximately the same population—if not you will need to repeat the PPS approach. As we will see in the caveat section, having final clusters of close to the same size limits some of the potential **biases** of the method. Regardless, if Hilandia happens to have 20 sub-divisions (such as wards, neighborhoods or quarters), they should be mapped out and the following steps should be followed to randomly select the final six.

Using PPS to select clusters from an area like Hilandia that has a large population follows the same steps as described for selecting clusters from the entire survey area.

1. Create a list of sub-divisions with their approximate populations.

2. Calculate the cumulative population for the sub-divisions. The last calculation will be the same as the total population of Hilandia.

3. Determine the sampling interval by dividing the total population of Hilandia by the number of clusters needed. In this example 6 clusters are needed from Hilandia.

4. Choose a random number to identify the starting point on the list to begin selecting clusters.

Beginning with the random number, use the sampling interval to identify the sub-divisions (clusters) of Hilandia that will be included in the survey.

Once you have selected clusters for the survey, it is important to visit those sample areas before data collection begins. Members of the Core Team should have already met with community leaders at the beginning of the pre-implementation phase to assess their needs and concerns and get community support for the survey. As a courtesy, field supervisors could visit the community leader in each sample area and let him or her know that the project and its local partners will be conducting interviews in their communities.

Community leaders can also provide useful information in terms of the layout of households within the sample area. It helps to draw maps of each sample area with the locations of each household. Survey teams can use this to at least select the starting household. In some cases the village leader knows where households with children 0-23 months live. These households can be listed. In this case the survey team supervisor can randomly select 10 of these households to interview. Ask the village leader to encourage the participation of selected households in the survey. Your local partners, who might already have a rapport with the target communities, can be very helpful in communicating with the target areas, particularly if your project is new to the geographic area.

You have now completed the first stage of sample selection, and we can now move on to describe the steps for selecting the households and respondents in the selected clusters. At this point, your supervisors and interviewers know where to go, but they now need a procedure for selecting households and respondents when they get to the community. Again, if you are entering a large community with multiple clusters, you will need to first randomly select which clusters to go to. We described above a procedure for doing this.

SELECTING HOUSEHOLDS (AND RESPONDENTS) WITHIN SELECTED CLUSTERS

In this section we will talk about methods for selecting the first household,[29] subsequent households and respondents. We will also discuss issues of how many households to include from each cluster; how many respondents to include from each household; and how to handle non-response.

We will focus on three methods of household selection:

1. Using a pre-existing household list and randomly selecting households before survey teams go to the field. This approach is mostly aligned with large survey procedures.
2. Mapping the households in a cluster and selecting households when the survey team is in the field.
3. "Spin the bottle technique" for selecting the starting point for household selection in the field. This approach is the most "rapid" of the three.

The core of the debate about which of these sampling approaches is best rests on the risk of erring from the rigor of appropriate random sampling principles. The first and second approaches listed above are considered to carry less risk than the third. There is, however, very limited empirical testing for what biases are induced by one versus another. And as we have stated repeatedly throughout this manual, non-sampling errors can represent a larger source of bias in survey implementation. The key message is to choose the approach you will be able to implement consistently and with rigor, bearing the purpose of your survey, and ensure the training and supervision required to do so.

We describe the implementation of these three approaches for each stage, starting with the selection of the first household.

[29] A household is a group of persons who share the same kitchen or hearth, or a group of persons who eat from the same cooking pot. You identify respondents from households.

A. Selecting the First Household

1) Preexisting household list

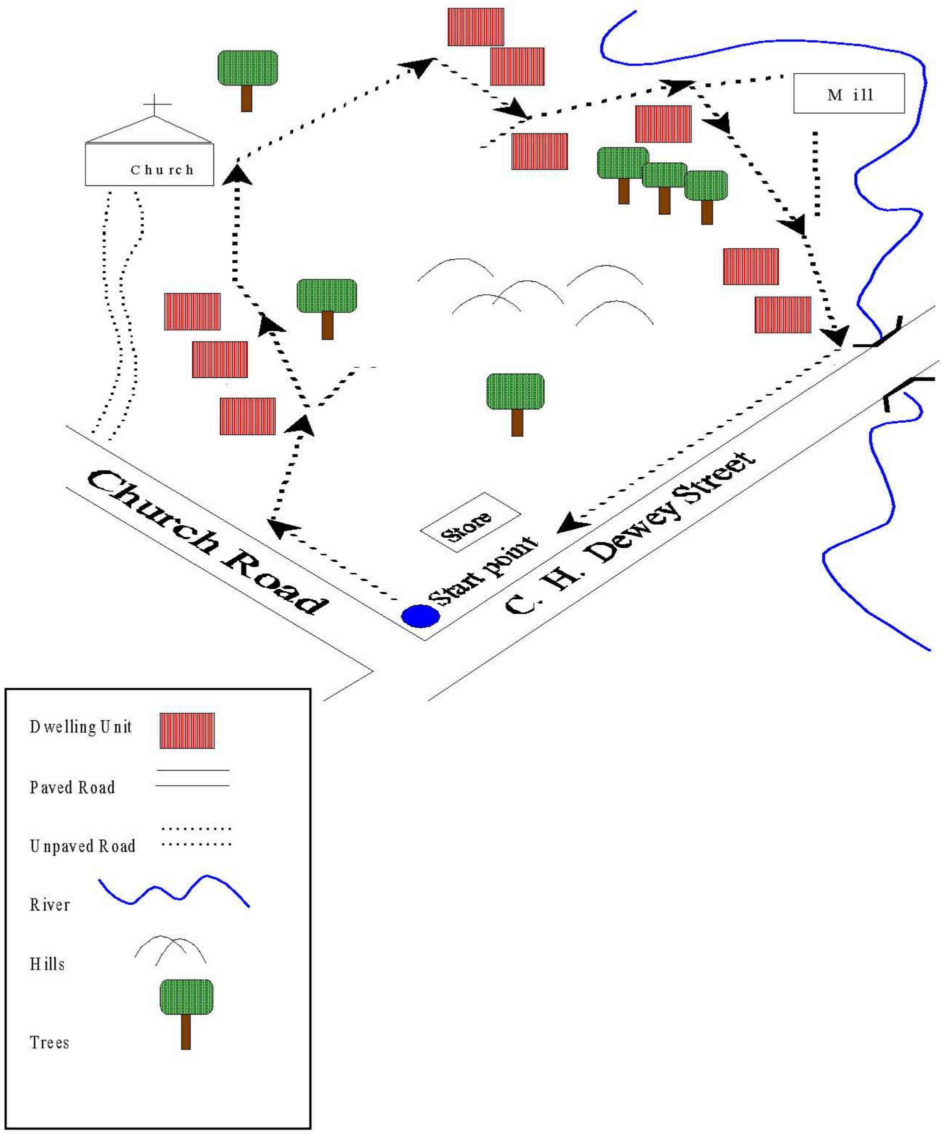
You can use this method, if a household list exists either from a national census or list created by an organization working in the survey area. For example, food distribution programs routinely make household lists in order to manage the distribution. You should make sure that this list is of good quality and is up-to-date. You also must be sure that survey teams will be able to locate selected households once they are in the field. In this case, random selection of households can be done as a desk exercise, up-front, before sending surveyors to the field. Annex 1 provides detailed references for this approach. In this case the first household selected corresponds to the first household on the list. In most cases where rapid surveys are implemented, a good quality household list is not available and there is not enough time or budget to do a census before conducting the survey.[30]

2) Mapping

This method is commonly used in rapid surveys, when a good quality preexisting list is not available. For this method the survey team arrives at a cluster location and works with key local people to map the location of each household. In some cases local people can even indicate which households have eligible respondents. Mapping works best for areas of 30 or fewer households. When a cluster is larger, it should be first divided into two to five sub-sectors of approximately equal size. One sub-sector should be chosen randomly. If this sub-sector still is too large, then divide it again until you reach a sub-sector with 30 or fewer households. Map this last sub-sector. Assign a number to each household in this sub-sector and randomly choose the first household. Maps should have clear boundaries (clearly delimiting the cluster—or in the second scenario below the subsection of the cluster). Maps should be comprehensive—listing all households, even those on the periphery. For this reason, the best way to sketch a map is to walk through the community with an informant and assure that you list all households (see Figure M2.1, drawn from the Malaria Indicator Survey (MIS) Supervisor Manual).

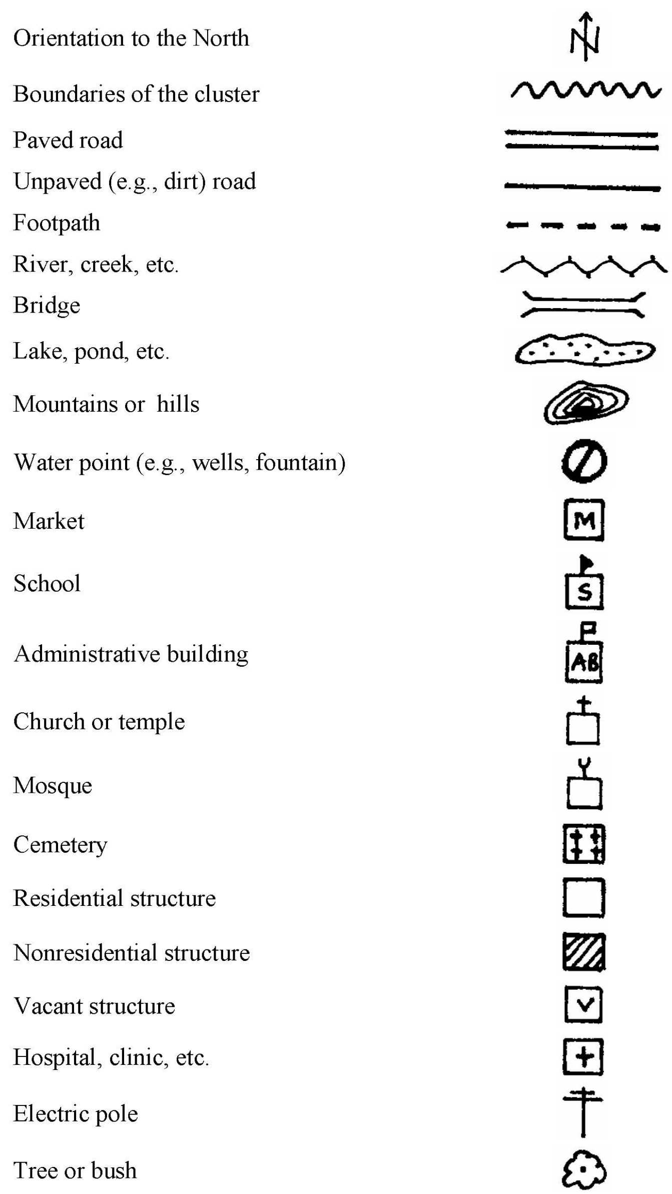
[30] While national surveys such as the DHS conform to this approach and use pre-existing household lists, it is important to remember that large amounts of resources and time are invested in these national surveys, which are normally only implemented every five years.

Figure M2.1: Example of a sketch map



This figure shows a progression around a cluster and indicates the importance of placing key landmarks on the sketch map so you or an interviewer can easily locate the household chosen. Use as many landmarks as you can to make it easier to identify households. The MIS Household Listing Manual provides an example of the kinds of symbols you can use on a sketch map to clarify landmarks (see Figure M2.2).

Figure M2.2: Examples of symbols for mapping



In one approach—recommended by the MIS (see Annex 1)—the listing of households based on mapping allows the random selection of households from the list. This encourages interview teams to work off well-developed maps to identify the first household and offer household listing forms to remedy the absence of a pre-made listing. This comes, however, at an "appreciable field cost," in terms of time in the field.[31]

[31] For details on the MIS approach, we encourage the reader to consult the MIS survey documents referenced in Annex 1: Key Resources.

3) Spin the bottle technique

This method can also be used when a quality household list does not exist. It is still frequently used in rapid surveys and is an accepted method, but it is the least statistically desirable of these three methods. There are however many instances where this is the most practical method to use. The following are the steps for this process:

1. Go to the population center of the sample area (the point in the community where the population is about equally distributed on all sides).
2. Select a smooth, level spot where you can place the bottle (or a ballpoint pen).
3. Spin the bottle.
4. When the bottle stops spinning, determine which direction the mouth of the bottle is pointing. The survey team should walk in a straight line in the direction that the bottle is pointing. (This step results in a random selection of which direction to follow.)
5. Count the households along this line.
6. When you come to the boundary of the cluster-stop. Let us assume you counted 20 households along the line.
7. Choose a random number from 1 to 20 to choose the first household. Go to that household, determine whether an appropriate sampling unit (respondent) is in the household and, if so, interview him or her.

Note: If two survey teams are conducting interviews in the same area, they should go in opposite directions.

B. Selecting Subsequent Households

1) Pre-existing household list

When a household list already exists, a random selection can be made of all households that the survey team will visit. This selection can be made ahead of time, and the survey team can be sent out with a predetermined list of which households to visit. This can help with quality control because supervisors can visit the preselected households to verify that interviews were conducted as reported by interviewers (see the caveat section below), and it provides an objective measure of non-response, a quality indicator for the survey.

2) Mapping

There are various ways to choose the next household with this technique. The most common method is to select the household with the "nearest door." This is the household with the nearest front door to the front door of the house where you are. If an eligible respondent is present, then you conduct the interview. If not, then you move on to the next household with the "nearest door." This procedure is followed until all the interviews are conducted. An improvement on the method is to visit every k^{th} household, for example every other household or every 4^{th} household.

While the next door approach is acceptable, potential bias might be introduced by having an interviewer select a "close" house that, for whatever reason (ease of access, condition of house, etc.), is not actually the closest to the door of house in which the first interview took place. For example, the next closest door could be across a stream, and the interviewer may decide to choose the closest door on the side of the stream where he or is.

This problem is eliminated when randomly selecting the households from a list before sending the surveyors in the community (as in the MIS). If you require the interviewer to use the next number on the list, you reduce the chance that he or she will wrongly select a convenient door. Forming a random list has the advantage of controlling for these factors. It does take effort and dedication to do this in the field and is therefore not normally the way subsequent households are selected, when the mapping technique is used.

3) Spin the bottle technique

With this technique, you either visit the household with the nearest door or visit every k^{th} household. If an eligible respondent is at the household, then you conduct the interview. If not, then you move to the next household until you have finished all the required interviews.

C. Selecting Respondents from a Household

Once you are at the first household (or subsequent ones), you have to determine what to do. Use Table M2.6 to guide your decision making process.

Table M2.6: Identifying respondents

If the type of respondent you are looking for:	Then:
Is at the household you selected	**Interview** that person **if** he or she consents.
Does **not** live at the household you selected	Go to the next-nearest household **from the front entrance** to the household you are at (or visit the next k^{th} household), and check at this next-nearest household. **Continue** this process until you find the respondent type you are looking for.
Lives at that household **but** is absent and far away (**more than** 60 minutes[32] away)	
Lives at that household, is absent **but** is nearby (**within** 60 minutes)	Go **find** the respondent with the help of a guide from the community. **If** you **cannot** find the person in the next 60 minutes, **go** to the next-nearest household **from the front entrance** of the household of the person you cannot find.

In rapid surveys, it is common practice to interview only one respondent per household if there are two or more eligible respondents. The rationale for this approach is that, in cluster designs, it reduces design effect (described below). If there are two or more eligible respondents in a household, you must use a random selection process to choose one. The process should be truly random and implies that interviewers have asked about whether there is more than one eligible respondent in the household.

The random selection process could be as simple as flipping a coin or randomly selecting a number between 1 and the number of eligible respondents in the household. It is critical that this

[32] See the Non-response section below.

be done randomly and that respondents not be permitted to self-select for the interview, as this could introduce a potential bias.

D. Non-response

For both rapid surveys and large-scale surveys, it is important to limit non-response. Surveys should be conducted at times of the year and times of the day when respondents are likely to be home. This means not conducting the survey at times of the year when people are away from home (such as planting or harvesting) and not conducting surveys at a time of day when people are not home. For example, if the survey is conducted in an urban area where most people work away from home during the day, then interviews should be conducted in the evening when respondents are more likely to be home.

Interviewers should be instructed to perform "call backs" to revisit households where the respondent is not home. Call backs can be repeated many times. A time limit can be set for these call backs.

The amount of time before calling back listed in Table M2.6 (60 minutes) is a suggestion—it could be considered a minimum. Also, in this table we recommend going and finding the person. This may not always be possible. In cluster sampling you may decide to go to the next eligible household and come back to the absentee household after having completed a successful interview. Calling back two or even three times is appropriate.

It is crucial to work with local leaders so that they can explain to community members the importance of participating in the survey. This will reduce the number of respondents who refuse to be interviewed. This works best when local leaders are included in the survey planning process and take ownership of the results.

Our take home message is that:

- Non-response is a threat to the validity of the survey.
- Aggressive plans (rigorous and systematic steps) should be made to encourage systematic call backs and to limit the risk of bias from non-response (see the Caveats and Potential Biases section below).

E. Summary Decision Algorithms

In the following pages, we summarize much of the information provided above in a set of "algorithms"[33] that you can use in training and that interview teams can use in the field. We would encourage you to adapt them to your context. They describe the following:

1. The process for selecting the first household in a cluster (Figure M2.3)
2. How to determine if you should conduct an interview in a household, based both on the type of physical structure of the household (something we have *not* discussed above) and the household composition (Figure M2.4)

---

[33] Diagrams are taken from the *KPC 2000+ Field Guide,* developed by the CORE Group and the Child Survival Technical Support project. See the reference for this document in Annex 1: Key Resources.

3. How to conduct the remaining interviews in a cluster, again accounting for the physical structure of the household (Figure M2.5)

The three figures provided here assume that your survey calls for interviews of caregivers of children under two years of age. Again, you will need to adapt them based on the respondent groups you are looking to interview.

Figure M2.3: Selecting the first household in a cluster

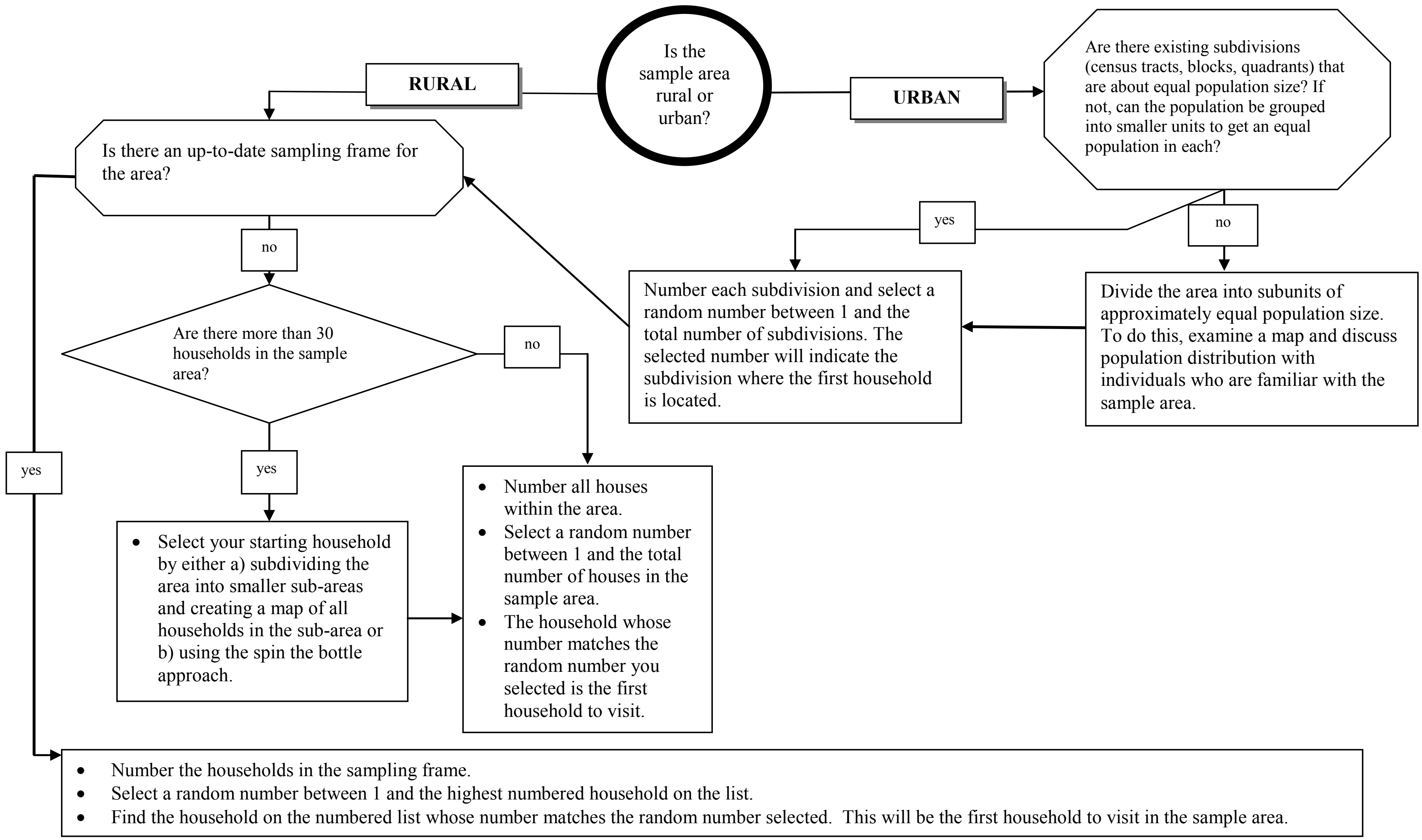


Figure M2.4: Choosing respondents based on the type of dwelling

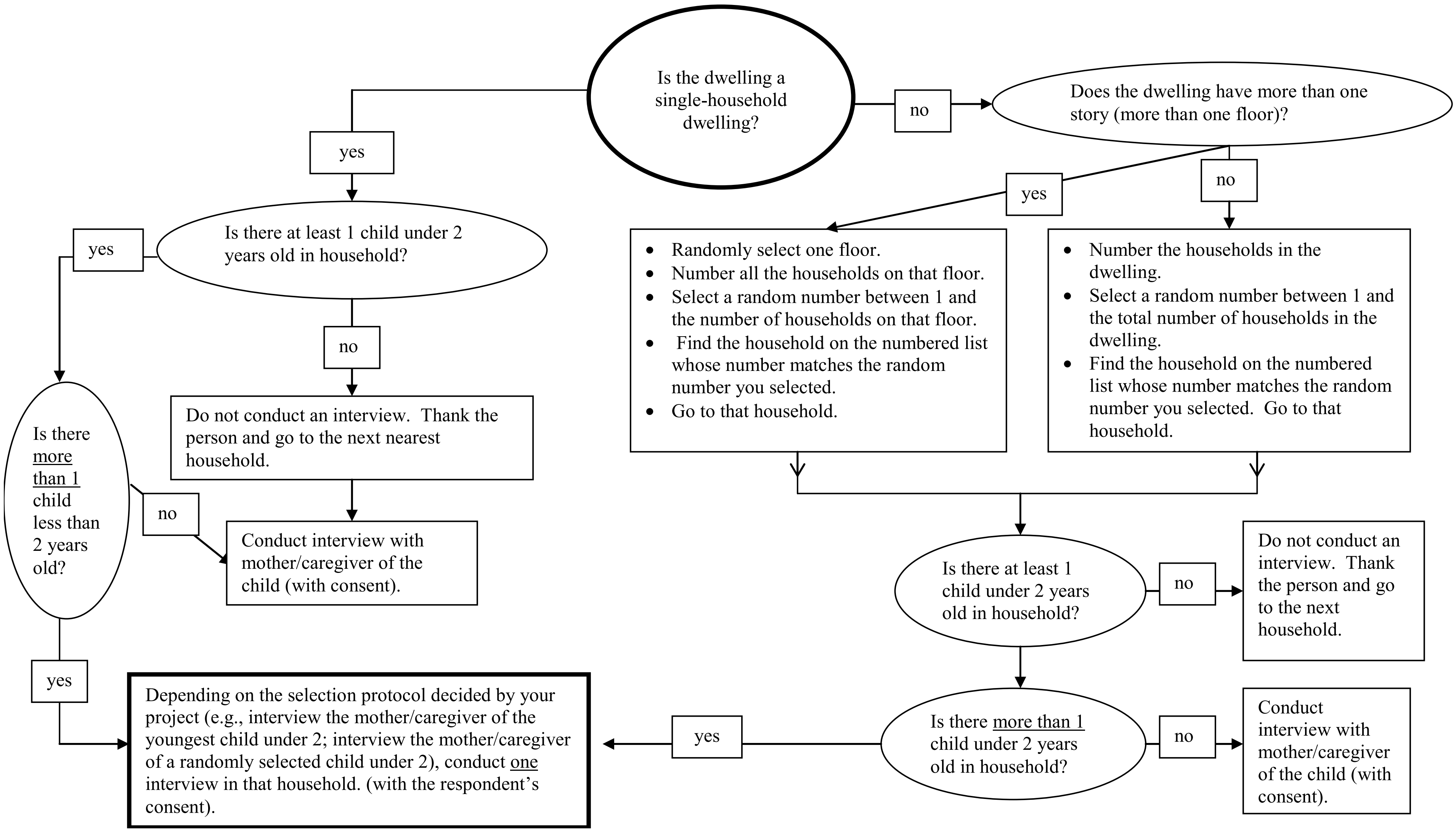
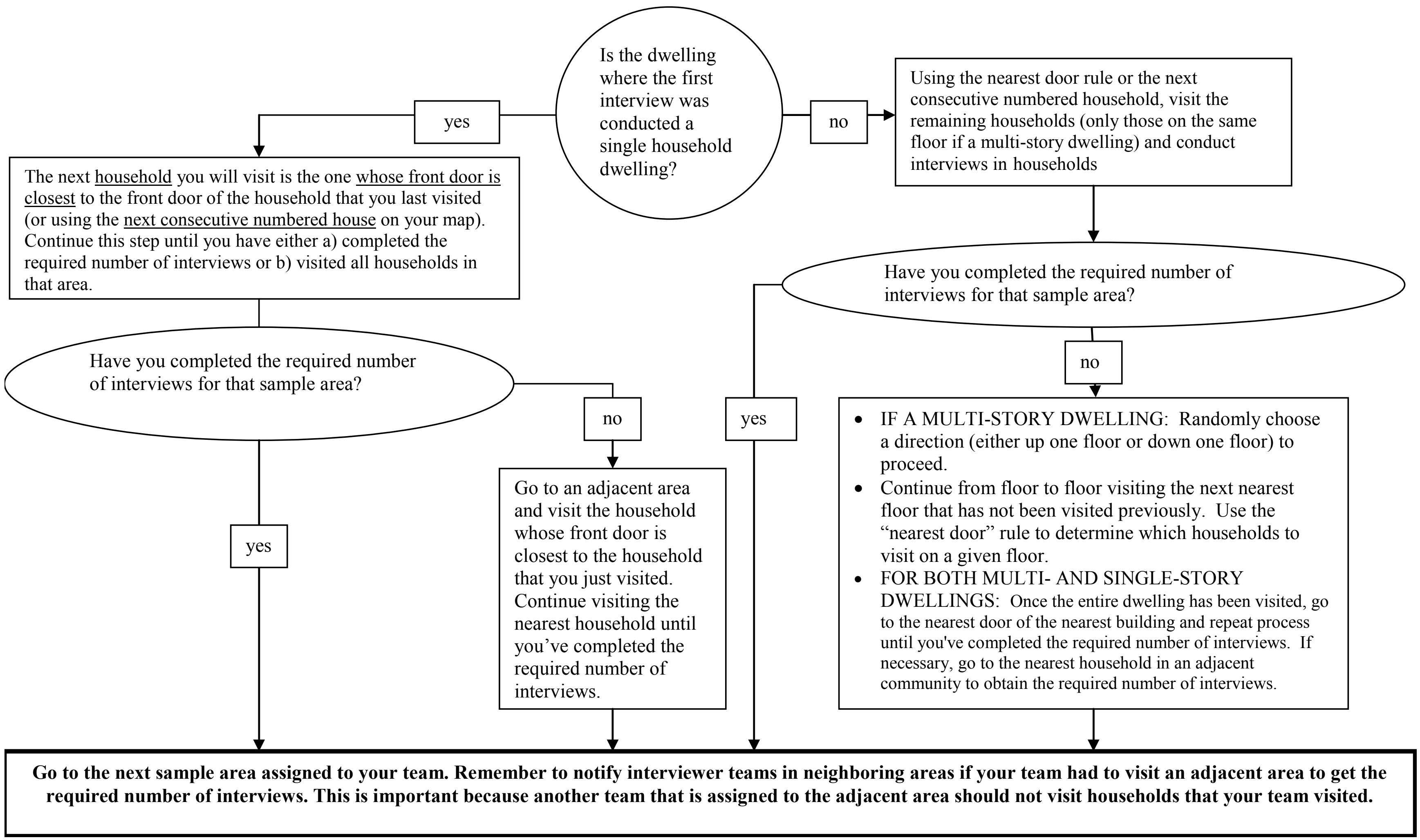


Figure M2.5: Conducting the remaining interviews in the cluster



We have now gone through step-by-step instructions for conducting a two-stage cluster sample—selecting first the clusters from a list of communities using systematic sampling from a list with probability of selection proportionate to size and then selecting households and respondents within clusters. Each situation is different in terms of the quality of information that is available to develop the sampling frame for the first stage and the many anomalies that may arise when selecting households and respondents.

The key is to do the best you can and remain steadfast in using random selection approaches when issues arise. Where you determine to do otherwise, you should state this in notes concerning a given cluster. Many situations have arisen over the years, such as a suddenly inaccessible part of a community (due to a downpour) or informants all being away from home due to market days or field work. Over time, you will develop approaches in your program area to deal with such issues, recognizing the potential biases they can introduce into a survey process. There is no ideal situation—each has its own challenges. Documenting decisions and committing to random selection processes is usually the best you can do.

Please review the instructions and examples for parallel sampling in the previous section if you have multiple questionnaires for which you may be able to interview different respondents in the same household. We will discuss parallel sampling in more detail in Module 3 in relation to its use in LQAS.

NOTE: The selection of the first household is a critical and sometimes time-consuming step in the rapid survey process. Consider having survey teams visit each sample area a day before starting interviews to correctly identify the starting household within each sample area. Please use the algorithms on the previous pages to assist you in decisions concerning cluster, household, and respondent selection. After these, we deal with the issues of limitations of cluster sampling and other more technical issues.

Some Caveats and Potential Biases in Selection Procedures for Households and/or Respondents

The procedures for selecting households and respondents described above represent tested and acceptable approaches to logistically challenging issues. However, while they *are* acceptable, they carry their own risk of bias and improvements or adjustments can be considered to reduce this risk. These adjustments, however, often take more time and effort to use and have not been widely used in rapid surveys until now.

We describe below some key points where large surveys, such as the DHS and MIS, are implemented differently than what is typical for rapid surveys and we provide brief explanations of how these changes help reduce biases, as well as the different set of constraints that come with those approaches. The decision to bring these approaches into rapid survey practice should be well-thought out. It is up to the reader to decide if it is worth the investment to implement these procedures. The following brief explanations should help the reader understand the issues. We encourage users who need more details about how to implement procedures used by DHS and MIS to consult the references for them in Annex 1: Key Resources.

The following table presents the main differences in the final selection of sampling units between large surveys such as the DHS and MIS and rapid surveys, followed by a discussion of potential biases rapid surveys need to pay attention to.

Rapid survey implementers must be aware of these risks, but they can take advantage of the comparative advantage of small surveys, which should make it logistically easier to maintain strict quality control through supervision than in larger surveys. Rigor, close supervision and attention to quality ultimately remain essential regardless of sampling approach.

Table M2.7: Traditional rapid surveys versus large surveys

Difference between traditional rapid surveys and large surveys	Potential bias
1. Pre-selection of households before interview teams are in the field vs. household selection at the time of interviews. Surveys such as the DHS and MIS carry out multi-stage sampling based on complete household listings (provided by the census). With a complete household listing, we can select a household sample by systematic sampling, in a central office or in the field. This will guarantee that every household has a real chance of being selected; and the interviewer receives the pre-selected list of households for the interview, which s/he will then simply use to conduct the interviews. In contrast, rapid surveys often perform household selection in the field either by creating a list on the spot or by using the "spin the bottle" technique.	**a) Bias that not all households will have an equal chance of being selected for the survey** When the households are selected in the field without an already prepared household list from a reliable source (such as census), it is hard to guarantee that every household has the same chance of being selected. For example, when using the on-the-spot mapping or "spin the bottle" methods, it may be hard to actually walk to the final boundary of the cluster/community, leading to some households on the periphery not being included in the listing. The problem in such a case is that households on the periphery of the cluster/community may be very different from those in the center. This is a fundamental problem with the on the spot mapping and "spin the bottle" methods. Unfortunately, failing to go to the periphery and include all houses is a real problem that has been encountered in the field. **b) Potential bias that only the easiest households will be selected for the survey** Since for large surveys, such as the DHS and MIS, there is a list of selected households, we have the option of checking to see if these households were actually visited. This possibility of quality control discourages non-sampling errors of interviewing the easiest to reach households. On the other hand, if no household listing is available, the interviewer has to select the households in the field. Because the household listing is done in the field, the interview team has to select the households to interview according to a predetermined procedure. You cannot assume that all the interview teams will follow (exactly) the instructions. Some will pick the easy-to reach households, or pass by those that are "not cooperative" or "not at home" quickly, in order to finish the interviews more quickly.

Difference between traditional rapid surveys and large surveys	Potential bias
2. Selecting a fixed number of *households* to visit in each cluster vs. collecting a fixed number of *interviews* (questionnaires) per cluster. Surveys such as the DHS and MIS select a fixed number of households to be visited per cluster. For example 15 households may be pre-selected to be visited. Once at the cluster site, the survey team visits all the households on the list; but they may only actually be able to perform interviews in 10 households. In contrast, in rapid surveys interview teams are normally given instructions to complete a fixed number of questionnaires, for example 10, in each cluster and to visit as many households as required to achieve this.	**c) Non-response bias** For large surveys, interviewers have a pre-determined list of households to visit in each cluster. As part of field procedures they record the result of the visit to each household on the list (i.e. interview completed, respondent not at home, no eligible respondent lives in household or refused to be interviewed). From this information a non-response rate can be calculated. A large non-response rate calls into the question the validity of the survey because those who do not respond may be different from those who were interviewed. If a non-response rate is calculated, survey managers can make better decisions about quality control of the survey or interpretation of results. In practice this results in varying numbers of interviews being conducted in each cluster and a need to use cluster weights when calculating estimates. (See UCLA reference in Annex 1: Key Resources for instructions on how to do this.) In rapid surveys, we only have anecdotal information from supervisors and interviewers about non-response. However, we have a fixed number of interviews in each cluster and can calculate estimates without using cluster weights (see below), thus reducing the time between when we begin a survey initiation and when we can start making decisions with the information. Both large surveys and rapid surveys have procedures for handling call-backs. It is important that these procedures are standardized and followed by survey teams. Quality control of how call-backs are handled is easier with a pre-determined list of selected households.

Difference between traditional rapid surveys and large surveys	Potential bias
3. Interviewing all eligible respondents from a household vs. randomly selecting only one respondent from a household Surveys such as the DHS and MIS interview all eligible respondents in a household. For example if the survey collects information about children 0-23 months and one household has one child who is 22 months and another who is 6 months, then information would be collected about both children. Another example is that if a survey targets women of reproductive age and a 42 year old mother and her 20 year old daughter live together, both would be interviewed. In contrast, in rapid surveys one respondent is randomly selected per household (i.e. one child 0-23 months or 1 woman of reproductive age).	**d) Bias of potentially adding an extra sampling stage** Since the interviewer may need to select one individual to interview (if there are more than two eligible respondents in the household), this adds an extra stage of sampling, and with it a marginal increase in the complexity of the sampling design. **e) Potential increase in non-sampling errors[34] of selectively choosing respondents with certain characteristics over others** If there is more than one eligible individual in the household because a mother and daughters live together (a *not* uncommon scenario), and if there are questions that the mother perceives to be sensitive, she may seek to "protect" her young daughters by saying either the daughter is not eligible, or not at home. So you may end up with a sample only having the mother or married women. This is just one example for how cultural factors can affect the validity of the sample. The reader will note that the selection of households against a list with the interview of all eligible respondents will also face some of these challenges, in terms of non-response, in terms of evasion, in terms of bias of response when an older or younger person is in the room. When conducting rapid surveys, it is important to think about how to minimize these potential biases. Regardless of the procedure implemented, survey managers must pay strict attention to quality control in the field. Supervisors and interviewers must be trained on the importance of actions, such as making sure not to avoid households from the periphery; informing the survey manager about issues of non-response, and the importance of thinking about randomness when making choices in the field. In addition, the survey manager must spend some time with each survey team to ensure that they are following strict procedures. Feasible procedures for call-back must be well thought out and respected.

[34]Recall that the estimates from a sample survey are affected by two types of errors: sampling errors and non-sampling errors. Sampling errors are the representative errors due to sampling of a small number of eligible units from the target population instead of including every eligible unit in the survey. Sampling errors are related to the sample size and the variability among the sampling units. Sampling errors can be statistically evaluated after the survey. Non-sampling errors—such as coverage errors, errors committed in survey implementation and data processing, etc. —are usually the most important sources of error. They are expensive to control and difficult to evaluate quantitatively. We can deal with them by improving supervision, standardizing all processes and reducing the likelihood that different interview teams will effectively use different procedures (for household selection, respondent selection, etc.) in the field.

Difference between traditional rapid surveys and large surveys	Potential bias
4. Applying sampling weights to each interviewed individual vs. assuming equal weights to calculate estimates. Surveys such as the DHS apply a sampling weight specially calculated for each interviewed household and as well as for each interviewed individual in their analysis. (A sampling weight is an inflation factor which extrapolates the sample to the target population, taking into consideration the probability of selection at each stage of the sampling design (cluster selection, household selection, and individual selection if the sample failed to include all individuals in the household). In general terms, sampling weights are used to make the sample more like the target population. The design weight also needs to be adjusted for household and individual non-response. In contrast, rapid surveys using 30x10 cluster sampling with primary sampling unit (PSU) selection proportional to size assume an overall equal probability of selection for each interviewed individual and almost never use sampling weights in their analysis.	**f) Potential bias of the estimates' confidence interval** Absent the application of sampling weights based on full household listing, rapid surveys face a risk of bias under two conditions: (1) if clusters represent PSUs of substantially different size, and (2) if the indicator values differ substantially across PSUs of different sizes. This potential bias cannot be totally eliminated, but is reduced by: • Sampling proportionate to size; • Mapping of communities into PSUs of comparable size (if PSUs are all of the same size and response rates are comparable the potential bias actually becomes negligible in PPS sampling); • Stratification along parameters likely to affect indicator estimates (such as rural/urban). In common practice for surveys such as EPI and KPC surveys, and with multi-stage sampling proportionate to size presented in this manual, the resulting bias has been assumed to be limited. Implementers should optimally take it into account when using full household listings and are referred to the DHS or KIS Sampling Manuals for details on sampling weight calculation. Absent this, even with a potential bias (of undefined direction) the estimates established under the 30x10 and LQAS approaches presented in this manual will allow establishing valid targets for local program officials who, by consistently applying the same sampling and estimating approaches, will be able to establish whether objectives have been achieved and (power permitting) whether change in indicators has been demonstrated.

SHORTCOMINGS OF CLUSTER SAMPLING

Despite its advantages, there are a few things to remember when deciding whether to use cluster sampling:

- Its inability to provide information on geographic subdivisions within the program area without LARGE increases in overall sample size
- The design effect

A cluster survey provides a project with coverage estimates for the entire program area. If you want coverage estimates for different management units (for example, supervision areas or health-facility catchment areas) in the program area, you will need to explore other sampling cause the sample sizes you will likely to need would be prohibitively costly using cluster sampling. As noted previously, LQAS may be a good option in such cases. We develop this methodology in Module 3.

Another important point to consider is that people of the same religion, socioeconomic status, or language/ethnic group tend to live closely together. Individuals who share these background traits are also likely to share similar behaviors and practices. This sameness (homogeneity) leads to an increase in the sampling error called the design effect (deff). The design effect exists because individuals selected from neighboring households within a given cluster are more likely to share the same knowledge and practices than individuals who are selected randomly from the general population. **To compensate for this bias, the size of a cluster sample should be approximately double the size of a simple random sample.** What this implies is that design effects (which vary by indicator) can be as much as two.

The problem is that one cannot calculate a sample's design effect for indicators until after the survey is over because its calculation is dependent on your results. This is why we use a "rule of thumb" and say that a design effect of two is a conservative estimate, requiring us to increase our sample size by two to achieve the same level of precision as a simple random sample. The Methodology and Sampling Issues for KPC Surveys manual already referenced refers to unpublished data from past KPC surveys, which provide the following estimates of design effect values encountered in different child survival projects in Nigeria, Indonesia, Honduras, Bangladesh and Papua New Guinea for a limited number of specific indicators:

Table M2.8: Design effect values

Indicator	Range of deff
ORT Use	1.06 – 1.79
Measles Immunization	1.06 – 2.16
Exclusive Breastfeeding of Children <6 months	1.08 – 1.91
Second Dose of Tetanus Toxoid in Pregnant Women	1.45 – 2.81

For more information on using software to calculate design effect from survey data please refer to the UCLA software training manual, the link to which you can find in Annex 1: Key Resources.

The 30x10 cluster sampling method leads to a sample size of 300 which, given the foregoing estimate of a design of 2, would be equivalent (again, subject to actual results) to a simple random sample of 150. Such a sample size gives us a precision of <10 percentage points for 95% **confidence intervals.** Generally, for project management purposes, this level of precision is more than acceptable.[35]

Let us say more about this point. Because a program is using a sample to estimate indicator levels, it should never regard an estimate from the survey as an absolute number. Each sample estimate falls within a range of possible values. This range is defined by the confidence limits. Confidence limits indicate the margin of error associated with a sample estimate. When we estimate a rate or proportion in a study, we expect the confidence interval to include the actual rate *most of the time*. How often is most of the time? Phrases like “95% confidence” give us an idea. When you calculate 95% confidence limits, you assume that there is a 95% chance that the actual coverage level is within the bounds of your estimated confidence limits. As noted, 30x10 cluster samples give us levels of precision for estimates of proportions, for example, that are <10 percentage points. This means that if we estimate that 30% of mothers appropriately feed their children based on our survey, then we would report that we are 95% confident that the true value for this behavior in the entire population is between 22% and 38%. The exact value depends on both our sample size and our design effect. Let us examine few examples to illustrate these points.

Here is the formula for calculating a 95% confidence interval for a proportion (percentage):

$$P = p \pm z \times \sqrt{(p \times q) \div (n \div deff)}$$

Where:

P	the actual rate/proportion in the general population
p	the survey estimated
q	1 – p
z	the confidence level (with a 95% confidence level, z = 1.96)
n	sample size
deff	design effect

Let us examine results from two surveys, the first in which the information was collected using simple random sampling (as noted, we rarely would use this approach—it is here as an example only). Recall that in simple random sampling there is no design effect, so the “deff” in the above formula is merely 1. The second survey used 30x10 cluster sampling and, for the indicator of interest, supposes that our computer analysis indicated that the design effect was 2.

[35] See Annex 3: Calculating Sample Sizes for Various Scenarios: Formulae and Examples on sampling for information on a more precise treatment of this issue. Note however that this estimate of precision assumes a design effect of 2. It is also based on the fact that the “widest” part of a confidence interval is for a coverage indicator at 50%. So if we were to have a design effect of 2 and found a coverage estimate of 50% we are saying that the precision would be <10 or narrower than 40-60%.

Table M2.9: Simple random sample versus 30x10 cluster sample

Simple Random Sample n=150	30x10 Cluster Sample n=300
Using a simple random sample, we find that 60 of 150 mothers of children 0-5m of age are exclusively breastfeeding. Therefore: P = the actual rate/proportion in the general population—this is what we want to describe. p = 60/150 = 0.40 q = 1 – p = 1 – 0.40 = 0.60 z = 1.96 n = 300 deff = There is no design effect so we set this to 1.0 $P = 0.4 \pm 1.96 \times \sqrt{(0.4 \times 0.6) \div (150 \div 1)}$ $\Leftrightarrow$ $P = 0.4 \pm 0.078$ So adding and subtracting this amount from 0.40, you get a range of [0.322 - 0.478]. If we want to express this result as a percentage, we move the decimal two places to the right: [32.2% - 47.8%]. Conclusion: *"There are an estimated 40% (CI: 32.2 – 47.8%) of mothers of children 0–5m who are exclusively breastfeeding in the program area."*	Using cluster sampling, we find that 120 of 300 mothers of children 0-5m of age are exclusively breastfeeding. Therefore: P = the actual rate/proportion in the general population—this is what we want to describe. p = 120/300 = 0.40 q = 1 – p = 1 – 0.40 = 0.60 z = 1.96 n = 300 deff = 2.0 $P = 0.4 \pm 1.96 \times \sqrt{(0.4 \times 0.6) \div (300 \div 2)}$ $\Leftrightarrow$ $P = 0.4 \pm 0.078$ So adding and subtracting this amount from 0.40, you get a range of [0.322 - 0.478]. If we want to express this result as a percentage, we move the decimal two places to the right: [32.2% - 47.8%]. Conclusion: *"There are an estimated 40% (CI: 32.2 – 47.8%) of mothers of children 0–5m who are exclusively breastfeeding in the program area."*

What this table confirms is that a 30x10 cluster sample with a design effect of 2 has the same precision as a simple random sample of 150. Of course, a design effect of 2 is a conservative estimate (design effects are likely to be lower in many cases). Again, for more information on using software to calculate design effect from survey data please refer to the UCLA software training manual, the link to which you can find in Annex 1: Key Resources.

SAMPLE SIZE ISSUES

We have talked a bit about the issue of sample size and its relation to precision of estimates and how you must consider design effects when considering sample size. In this section, we provide more information on sample size determination. All along we have suggested that a cluster sample size of 300 (30x10) has been shown to be a useful size for most management decisions that program staff need to make. Even with a design effect of 2, it provides levels of precision <10 percentage points.

A sample size of at least 300 is also usually adequate for looking at sub-samples (such as children 0–5 months to assess exclusive breastfeeding), because it usually guarantees enough respondents in the sub-sample to keep precision levels at around 10%.

Despite these considerations, different programs may need to increase sample sizes to increase precision. In the next section we talk about options for improving on the 30x10 cluster design, but before we discuss them let us be clear about the reasons why we say that a sample size of 300 is typically adequate and then provide information you can use should you choose to adjust sample sizes.

In general, rapid surveys are most useful (beyond their use at baseline to help identify priorities) to help programs answer the question "Have the program objectives been reached?" This means that rapid surveys rely on a single sample taken at a point in time (cross-sectionally), the results of which are compared to targets. We have already seen how to calculate confidence intervals for point estimates coming from a survey. If the confidence interval's values are greater than the objective, then we can say that we have evidence that the program has contributed to reaching the targets set.

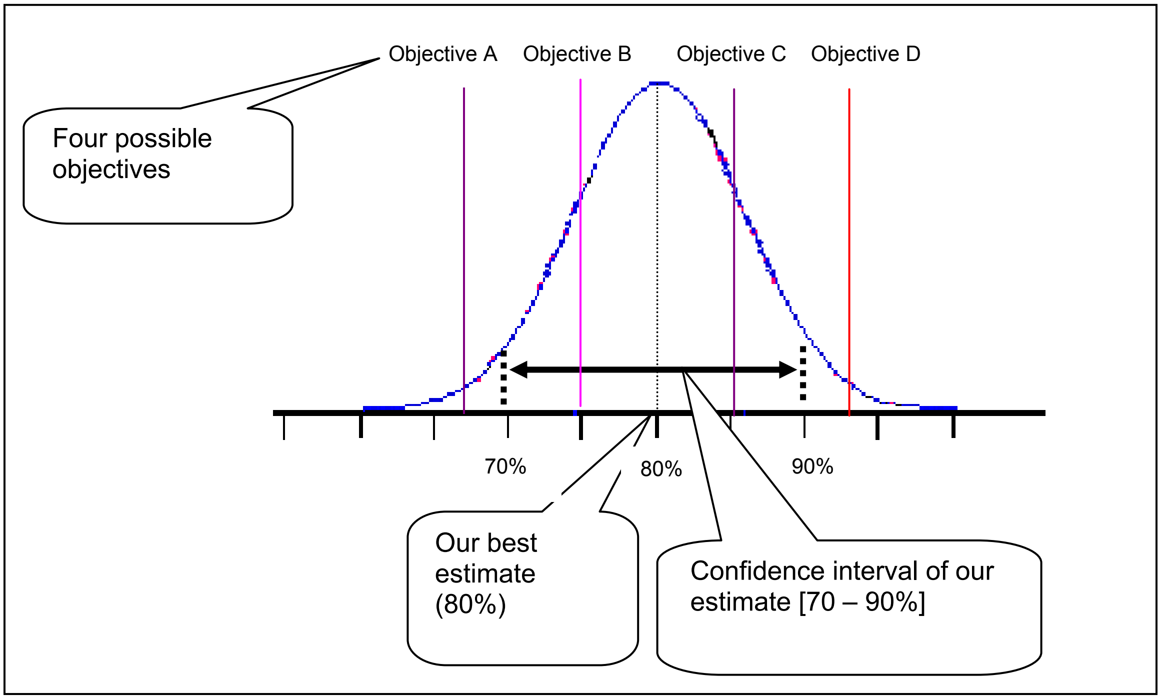
Using the example from the table above, suppose we had set a target that, by halfway into the program intervention, 30% of mothers of children 0–5 months would be exclusively breastfeeding these children. We may have set that target because at baseline we learned that only about 10 percent of mothers were doing so. If we used the results from above survey taken about halfway through the program, we would be able to conclude that we had evidence that in fact we had reached that target. This is an example of the most straightforward use of rapid surveys: one sample compared with a target. Notice that we need only calculate a point estimate and its attendant confidence interval to be able to reach our conclusion. We can then choose our sample size based on how much precision we want in our estimate. Because the estimate is compared with a fixed value (the program performance objective), only one level of imprecision has to be taken into account.

Let us examine this point in more detail with another example. Let us consider that, based on our sample of 300 survey respondents; we have obtained an estimate of 80% for the coverage rate for a given indicator. The meaning of choosing a confidence level of 95% and achieving a precision level of ±10% is that: The true coverage rate in the entire population (which we do not know but estimate at 80%) has a 95% chance of being within 10% of our estimate, (in this case between 70% and 90%).

Let us now consider four possible levels of objectives that our project could have initially set out to reach using Figure M2.6 below:

- Objective A (e.g., 68%) is outside and below the confidence interval of our estimate.
- Objective B (e.g., 75%) is below our estimate but within its confidence interval.
- Objective C (e.g., 85%) is above our estimate but within its confidence interval.
- Objective D (e.g., 93%) is outside and above the confidence interval of our estimate.

Figure M2.6: Comparison of a final estimate to a pre-set objective (hypothesized at four different values)



Source: Methodology and Sampling Issues for KPC Surveys

Table M2.10 suggests the conclusions that can be made, from a statistical perspective and from a management perspective about these four different situations. Notice that our results are stated as hypotheses: we are essentially hypothesizing that we have reached the objective we set. We use the confidence interval (coming from our sample results) to state whether we have evidence for this hypothesis or not.

Table M2.10: Assessing achievement of objectives

Level of objective	Statistical conclusion	Managerial conclusion
Objective A	The difference between our best estimate and objective A is statistically significant. We reject the hypothesis that the population coverage rate is equal to the pre-set objective. We are more than 95% confident to have reached our objective.	Our objective has been reached. → Continue activities, or → Expand program, or → Plan transfer and phase-out → etc.
Objective B	We cannot reject the hypothesis that the population coverage rate is equal to the pre-set objective. Our best estimate is that the population coverage rate is 5% higher than objective B, but the difference is not statistically significant.	Our objective has probably been achieved (our best estimate is that we are 5% above objective). There is no evidence that we have failed to reach our objective, but we cannot prove that we have achieved coverage higher than the objective. → Was performance homogenous in all local areas of interventions? → Do other sources of information support or contradict our conclusion?
Objective C	We cannot reject the hypothesis that the population coverage rate is equal to the pre-set objective. Our best estimate is that the population coverage rate is 5% lower than objective C, but the difference is not statistically significant.	Our objective has probably not been achieved (our best estimate is that we are 5% below objective). There is no statistical evidence that we have failed to reach our objective, but we cannot prove that we have achieved coverage higher than the objective. → Was the objective too ambitious? → Was performance homogenous in all local areas of interventions? → Do other sources of information support or contradict our conclusion?
Objective D	The difference between our best estimate and objective D is statistically significant. We reject the hypothesis that the population coverage rate is equal to the pre-set objective. We are more than 95% confident that our program has not reached its objective.	Our objective has not been reached. → Was the objective too ambitious? → Was low performance homogenous in all local areas of interventions? or → Are specific local areas responsible for overall low performance? → Plan and implement corrective measures

Source: Methodology and Sampling Issues for KPC Surveys

As illustrated in this theoretical example, once a hypothesis is formulated (e.g., an estimate is equal to or superior-or-equal to a set value), statistical tests achieve a high level of certainty when they can disprove this hypothesis.

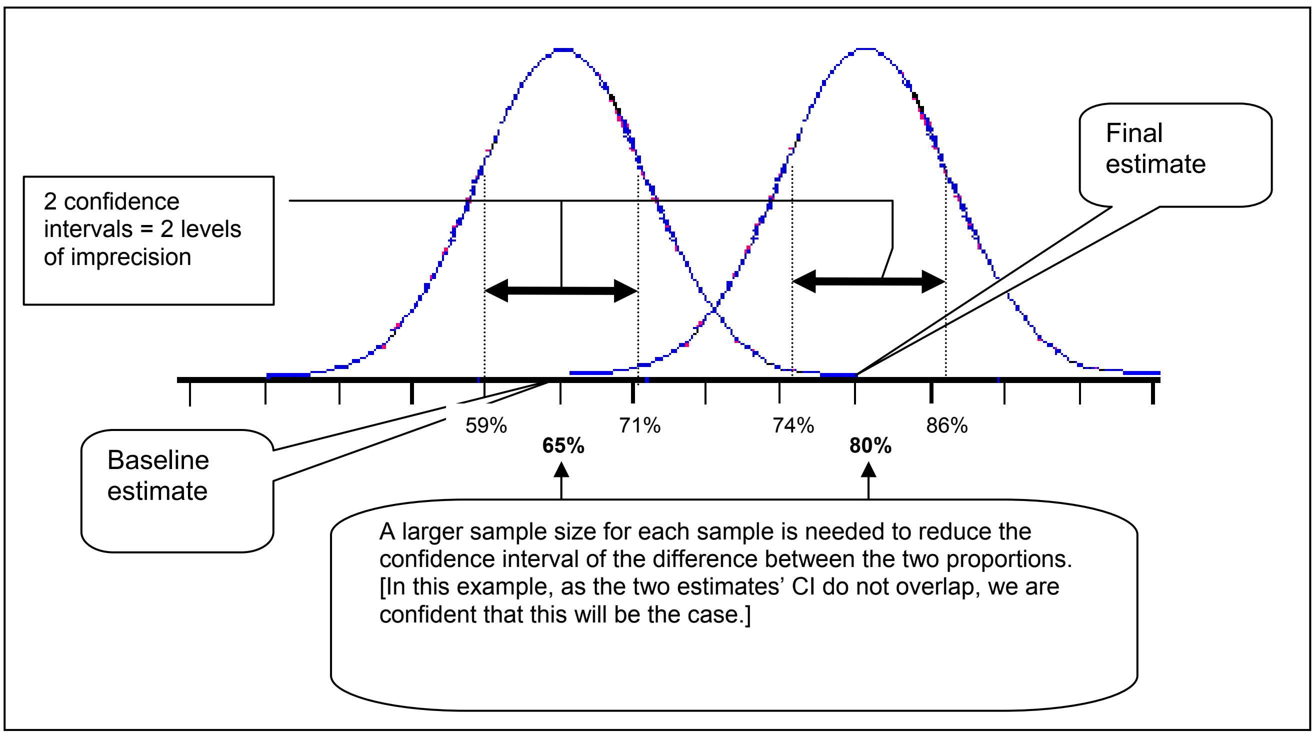
Defining objectives and assessing achievement of objectives belong in the field of management and decision-making. Relatively small samples are allowed, and field staff can collect and compute the data. The 30-cluster sampling scheme for the rapid health survey was chosen to achieve this level of precision and reliably answer these management questions.

The story is quite different when we move from asking whether we have reached a pre-set target to asking the question: "Has change taken place in the intervention population for selected indicators?" Intuitively this makes sense because it implies that we have taken two samples (one at the beginning and one at a second time), and for each sample we must calculate confidence intervals. That is we now have two estimates from samples and the attendant uncertainty twice over.

In other words, comparing a final estimate with a set objective only introduces one level of imprecision: that of the estimate. In order to compare final and baseline estimates, we have to take into account two levels of imprecision, those of the baseline estimate and the final estimate, and thus we require an increased sample size *for the two* surveys.

While in the previous example, we were testing whether a fixed value was below or within the confidence interval for our estimate, we are now testing whether two estimates, each with its own imprecision, are statistically different. In the first case, we established with 95% confidence that the objective of 68% coverage (objective A) was not in the confidence interval of our estimate (70% to 90%). In this second situation, we must calculate a confidence interval for the difference between the two samples. (We can also test whether this difference is significantly different from zero, the two methods being equivalent.) In order for us to statistically demonstrate a difference between the two estimates, we will need a larger sample at each phase, thus reducing the two levels of imprecision.

Figure M2.7: Comparing estimates from two samples



Source: Methodology and Sampling Issues for KPC Surveys

The decision to make a before-after comparison must be taken before the onset of the program, and the sample size of the baseline survey must be calculated appropriately. (Correcting the final survey sample size for a late decision to "power the survey" in such a way as to be able to make a comparison is sometimes possible but not entirely recommendable.) Although the desire of program managers to compare the final and baseline estimates for different indicators is understandable, they should only do so under the following conditions:

- First, establish whether the confidence interval for the final estimate excludes the program objective for a given indicator. Statistical evidence should be used when it exists. This first level of assessment is important in establishing with what level of certainty the program is thought to have reached its objectives.
- Then, if a comparison of baseline and final results is presented, the evaluator should make explicit whether the sample size for the two surveys did or did not attempt to be able to demonstrate a difference with statistical significance.
- An observed difference should be reported with its confidence interval.
- Finally, inasmuch as possible, other sources of information should be used to try and assess whether the observed difference may or may not be genuine.

We have attempted to show here the implications of using the results of a rapid survey to assess evidence for achievement of targets versus using them to compare change in indicators over the life of a program. The most important message is that, in most cases, a larger sample size in each survey will be required to assess change in indicators. We provide the appropriate formulas for calculating samples sizes in Annex 3: Calculating Sample Sizes for Various Scenarios: Formulas and Examples. This annex provides sample size formulas and examples of their use for different survey purposes (comparing results to a fixed target, comparing final evaluation results to baseline, etc.).

One final issue related to sample size calculations concerns assuring you have a large enough sample of "sub-groups" of interest in your population to calculate estimates with adequate precision. We have already discussed one way to deal with this by using parallel sampling. In addition to this approach, the following table, adapted from Methodology and Sampling Issues for KPC Surveys (the link for which is in Annex 1: Key Resources), provides other options—and the strengths and limits of each—for assuring you have enough respondents for sub-groups of interest in your survey.

Table M2.11: Methods for calculating sample size

Method	Description	Advantages	Constraints
Stratifying	The population is divided in as many strata as specific subgroups of interest (sex, age-groups). The sample is predetermined to include a defined number in each stratum, with the same contribution from each cluster.	The number of respondents in each stratum is known beforehand, the level of precision of the survey is improved (by decreasing the deff), and no group is under-represented.	Adds complexity to the sampling strategy.
Increasing the entire sample size	The entire sample size is increased, and it is expected that the sample size of each sub-group of interest will also be increased proportionally. The number of clusters can be increased, or the number of respondents within clusters can be increased.	Simple and straightforward. Precision will be improved by an increase in the number of clusters, as opposed to an increase in cluster size.	Increasing the number of clusters is costly. Increasing the size of the clusters will increase the deff. Leaves to chance the selection of a sufficient number of children in a given age group.
Over-sampling in the group of interest.	To have a precise coverage rate for immunizations among 12-23 months old children, the surveyors will be instructed to survey 10 children in each cluster, as they would normally, but then to interview additional mothers, exclusively about the EPI questions, until 7 children in total have been surveyed in the 12-23 months old age group. (In this case, this strategy would achieve a sample size of 210 children for the EPI questions). In most instances, this will only require adding two more children in the cluster. These two children will only be included in the analysis of the EPI questions.	Since these additional two are surveyed exclusively on the questions of immunization coverage, the added time (and cost) is relatively limited.	
Parallel sampling	A specific survey is administered for two different age groups. The sample size is calculated for each one separately, and two different questionnaires are prepared. The mothers are sampled from the same clusters and the same households and the two surveys are conducted through the same surveyors, using the same logistics. For each group, the desired level of precision is chosen and determines the size of the cluster for the age group.	Cost-efficient use of logistical resources to obtain a predetermined level of precision in two different groups. Similar to stratification.	Requires multiple questionnaires.

IMPROVING ON THE TRADITIONAL 30X10-CLUSTER DESIGN

In addition to these sample size considerations, if your program wants to generate more precise and/or programmatically meaningful data, it can make changes to the cluster design or choose another methodology altogether.

Because precision increases with increasing sampling size and smaller design effects, you can try to increase precision by dealing with one or both of these issues. First, you can increase the sample size of a cluster survey in one of two ways: (1) increase the number of clusters, (2) increase the number of interviews conducted in each cluster.

Increasing the number of clusters is much more desirable than increasing the number of interviews in each cluster. That is due to one factor: the design effect. As you increase the number of interviews in each cluster, you are likely to include more people who are alike in terms of the knowledge and practices in which your project is interested. Therefore, increasing the number of interviews in each cluster could merely increase the design effect. This results in survey estimates that are less precise.

To deal with the design effect problem, you could try to reduce the homogeneity within each cluster. One way to do this is to change the protocol for sampling households. For example, instead of instructing interviewers to visit the nearest household, your project could instruct interviewers to visit the third closest or fifth closest household. If you sample households farther apart, they might be less likely to share the same knowledge and practices. Therefore, by using the third closest or fifth closest rule, your interviewers are visiting a broader area of the community. This could help to reduce the design effect, though, again, you cannot know that until the data have been collected and design effects calculated.

Reducing homogeneity within each cluster might improve the precision of your survey's estimates. However, you should keep in mind that it might also take interviewers longer to complete the 10 interviews required in each cluster. This is especially true if populations are dispersed (households are located far apart from each other). Once again, you will have to decide whether small improvements in precision are worth the extra time and resources.

FINAL THOUGHTS

If you desire more information on sample size issues please consult Annex 3: Calculating Sample Sizes for Various Scenarios: Formulas and Examples. We have attempted here to provide an intuitive understanding of cluster sampling and have provided a step-by-step example that we believe you can follow to develop your own sample. For those who desire a more sophisticated treatment of this information, along with Excel programs that enable you to specify different levels of precision and or sample sizes, please refer to the documents referenced in Annex 1: Key Resources.

Before moving on to LQAS, we note that the selection process for interview sites (not clusters) using LQAS is the same as for two-stage cluster sampling. Indeed, all of the procedures and steps outlined above are used to select interview sites and respondents in LQAS as well. Therefore, we will not repeat them below. We will note some conceptual differences between cluster and LQA samples but spend more time describing the uses and interpretation of LQAS results.

MODULE 3: PLANNING AND IMPLEMENTATION OF RAPID SURVEYS WITH LOT QUALITY ASSURANCE SAMPLING METHODOLOGIES

UNDERSTANDING LQAS

Even for regular users of LQAS, delving into its statistical idiosyncrasies is challenging, and practitioners often end up correcting each other about the validity of a particular inferential statement because LQAS requires us to make articulate inferences in ways that we are not used to articulating them. We cannot deny this. On the other hand, LQAS is a valid sampling approach for rapid surveys, and for managers it offers action-oriented information, which has made it popular among the many who have learned to apply it.

We therefore need to start addressing LQAS from a manager's perspective. Once the fundamentals are clear, we will progress into the statistical specifics of the method. We hope the reader will feel confident about the method and ready to use it after reading this module, if it serves his or her management purpose.

WHY USE LQAS—AN EXAMPLE

Before going into details about LQAS, let us describe a situation for which its use might be appropriate and examine what it can and cannot tell us. Suppose you are the manager of a health project that is being implemented in a number of different geographical areas, with different staff responsible for the implementation in each area. Suppose, for the sake of discussion, that you have five different zones where you are implementing your program (call them supervision areas—see Figure M3.1) with five different supervisors responsible, each for one zone. Your job is to make sure that all of them are making progress toward the objectives you have set. It would be useful to have, for each supervision area, information allowing you to say whether it appears that the zone is achieving or is close to achieving the objectives, or whether a particular zone is falling far below expectations. If you had this information, you could use your scarce time resources to support the zone that is falling far behind. Your rationale is that even if a given zone is not quite keeping up with targets, you need to identify those that are falling substantially behind.

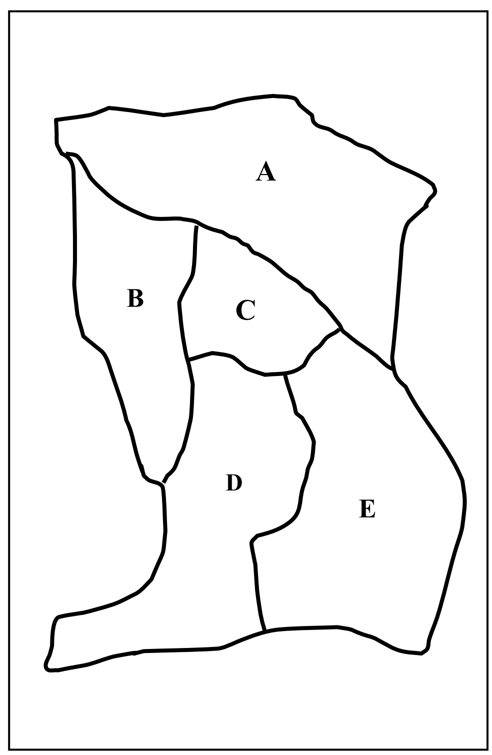


Figure M3.1: Supervision areas

Together, A, B, C, D, and E represent the program area. A, B, C, D, and E represent five supervision areas.

A rapid survey sampling method that allows you, your staff, or a survey partner to take a relatively small sample in each area (using the random sampling approaches already described above for selecting communities, households, and respondents) and determine, yes or no, whether each zone was close to or already reaching targets or falling far behind would be useful. The key is to take a small sample that allows you to make this yes or no judgment for each SA. This, in essence, is what LQAS is designed to enable you to do.

LQAS allows you to take a small sample in each zone (for now, accept without proof that this a sample size of 19 is sufficient) and use that information to make the yes/no judgment for each zone. You can train staff in each zone in the sampling methods previously described, and they can routinely select a sample during a specified time and interview respondents concerning key indicators. Since there are only 19 individual respondents, staff can quickly tabulate results and, using a simple predetermined rule (described below), can determine whether the program in that zone is reaching the target or falling below the target and in need of attention. In some cases the program may decide, at a given time during implementation, to focus on only a subset of program indicators and make this judgment for each indicator in this smaller selection rather than conducting a full survey of all indicators. Used in this way LQAS is a very rapid ongoing progress tracking (monitoring) tool that program staff can use on a regular basis.

There is one more thing LQAS can achieve, and this is why it is discussed as a survey sampling approach (and not just a quality control statistical method—the area for which it was first developed). In addition to allowing a yes/no judgment concerning each of the five zones in our example, the results of the five samples can be combined to provide an estimate value for the selected indicators for your entire program area (provided the combined sample is large enough, as we will see later).

Of course, as with any sampling approach, LQAS comes with a measure of imprecision or error. There is, specifically, potential for misclassification—either indicating that a zone has reached a target when, in fact, it has not, or indicating that a zone has not reached a target when, in fact, it has. This corresponds to the important statistical concept of "risk," we will discuss it below. More importantly, please note that what is not possible with LQAS is to obtain a valid indicator estimate for an individual zone. Once again, LQAS merely allows you to determine whether there is evidence that a given zone is reaching the target or not. This is the key feature of LQAS and the main reason for its use.

How LQAS Works—A Closer Look and Another Example

With this explanation in mind, let us look more closely at how LQAS works. We have already described the use of LQAS for monitoring progress; it can also be used for larger, more inclusive baseline, mid-term, or final evaluations. This is described in more detail below. As noted, the indicators, questions, and community, household, and respondent selection processes are the same for LQAS as for 30-cluster surveys. LQAS is not, however, a cluster sampling approach (see the text box below). Consequently, rather than selecting a cluster and then interviewing multiple people in that location, in LQAS only one respondent is taken per interview site selected (there are some exceptions to this, which we will discuss).

What then, besides this important last point sets LQAS apart from the 30-cluster sample? First of all, in each zone a very small sample is taken. For cost efficiency reasons, a sample of 19 is often chosen, so we will assume that this is the case here (we will see later how it can vary). If, as in the example we began to develop above, there are five supervision areas (SAs), this would lead to a total sample of 95 (19*5=95). Each SA is responsible for randomly collecting 19 responses for each indicator of interest. The natural question is: What can we and what can't we say about the SA given this small sample size? Let's use a simple example to answer this question.

A supervision area is not a cluster

In LQAS, you randomly select 19 individuals from each zone or supervision area. In cluster sampling, these supervision areas are not considered at all—instead, clusters within the overall survey area are randomly selected; then individuals are selected based on being within a selected cluster.

Suppose that we are implementing an HIV/AIDS prevention program among women age 15-49 along with some child health activities focusing on these women. We may decide that it is important for all women in the program area to know at least two ways that HIV spreads from person to person. One indicator of interest would be the percentage of women 15-49 (in the program area, consisting of all SAs) who can name, in response to a survey question, at least two ways that HIV spreads (or how to prevent such a spread). Further, based on prior surveys and the amount of effort we are giving to this issue, we hope to have at least 80% of women in the program area know at least two ways that HIV spreads by the time the program ends. This is our coverage target.

We now randomly select (using methods described in the cluster sampling section above) 19 women in each SA, interview them, and for each one we give a score: "0," if they do not know at least two ways and "1," if they do. We then add up all the "1s" and get a total score (out of 19) for the SA in relation to this indicator. Suppose that we find that 14 women (out of 19) in an SA know at least two ways that HIV spreads. What can we conclude? (IMPORTANT: We will not even think about producing a percentage estimate for this SA, based on this proportion of 14 out of 19. This sample is too small to produce a valid estimate.) But do we have evidence that this SA reached the target of 80%?

To answer this question it is best to think about it this way: *If, in reality, 80% of women in this SA know two ways that HIV spreads, how many women out of 19 in our survey sample should be able to name at least two ways?*[36] In other words, how many women should we count as "1" out of 19 in order to conclude that this SA has reached the target? The answer is that we should find 13 out of 19. If we get 13 or more who know two ways to prevent the spread of HIV, we will conclude that this SA has reached the target (80%) for this indicator. The number 13, as we will see later, is what is called a "decision rule[37]." In our example, we would say that we have evidence, since 14 out of 19 women knew of two ways for HIV to spread, that this SA can be *classified* as having reached the target meaning that we will not consider it in need of a remedial intervention.

[36] Obviously, we do not know whether 80% in this SA know two ways that HIV spreads. If we knew that we would not have to take the sample.

[37] While we are presenting only the decision rule for 80% here, decision rules for all percentages (at 5% intervals) are available and presented below.

Notice how we said this: We will *classify it as if it has reached the target.* Unfortunately we don't know if it has reached the target or not because with a result of 14 (in this example) it *is* possible that what we actually have is a supervision area that is close to 80% but has not reached 80%. In other words, with LQAS we may misclassify SAs that are close to the target but have not reached it by saying they *have* reached it. So, we should be cautious and not conclude that we have reached 80%. However, there is very little chance that we have misclassified an SA that is far below the target with "far below" being 50%. So by saying we *classify it as if it has reached the target* we are saying that we are not worried that it is falling far below the target and in need of help. This is explained in greater detail below.

So, we must ask the question: "How sure can we be of having reached the right decision about our SA?"

Here is what LQAS tells us:

- If the population of the SA is truly one in which 80% of women know of two ways that HIV spreads, there is a greater than 90% chance that we will find 13 or more in our sample of 19.[38]

There are a lot of numbers to follow in this statement, but it boils down to this:

- If the population of the SA is truly an "80 percent knowledgeable population", we will find 13 or more knowledgeable women out of the19 interviewed almost all the time (with "almost all" meaning more than 90%). Consequently, when we find 13 out of 19 women in a given SA knowing at least two ways, we will conclude that this SA has reached the target.

There is a chance or risk (less than 10%) that even though the population is an 80 percent population we will find less than 13 who know two ways in our sample of 19. In this case, we would conclude—wrongly—that the SA did not reach the target.[39] With a sample of 19, there is a small chance (less than 10%) that we will misclassify the SA saying that it did *not* reach the target when *in fact it did*.

The problem is better understood from a management perspective: imagine that you are the supervisor of an SA and that—due to the hard work of your team—you have actually reached the target. You might get quite discouraged if your SA is classified as not having met the target following the LQAS decision rule. Any sampling approach carries the problem of sampling error and this is a risk we must live with. LQAS is set to minimize this risk (10%), but its importance is worth remembering. However, if the SA is using LQAS to routinely assess progress (for example every couple of months), the risk that an 80 percent population would be misclassified twice becomes so small that it is unlikely to happen.

---

[38] This statement, and all conclusions like it using LQAS, is based on the cumulative binomial distribution of probabilities. In this case we know that the probability of finding 13 OR MORE who know two ways to prevent the spread of HIV in a sample of 19—from a population in which the true percent who know two ways is 80%--is actually 93%.

[39] This is called an alpha error. See the following sections for more information and the section below "LQAS Applications: Special Topics" for a detailed description.

MORE ON RISK OF MISCLASSIFICATION

You might be thinking: "Well look, why don't we merely increase the sample size to say 20 or 25, that is still not a very big sample and that would reduce the chance of misclassification (wouldn't it?)."

While logical, it turns out that increasing sample size even by a great deal does not reduce by very much the chance of misclassifying an 80 percent population as not being an 80 percent population. And this explains why we use a sample size of 19. With this sample size, and for all possible percentages, the probability of misclassifying a successful SA as not successful is less than 10%. Generally we are willing to live with that chance to make management decisions because we will RARELY misclassify a "successful" SA as "unsuccessful".

But what about the risk of concluding that an SA has reached the target when in fact it has not that we alluded to above?[40] Let's formalize our understanding of this since the story is a little different for this question.

LQAS is set up so that it does not discourage workers by saying they did not reach the target when in fact they did. As a result, and because it relies on a small sample size, it is not as good at accurately classifying those who are close but have not reached the target. Recall what we said earlier, however—what we really want as managers is a way to identify SAs that are falling far short of the target. Those that are falling short by a small margin may be misclassified, but we are less concerned about them. We want to identify those that are falling far behind. So what is the chance that we will misclassify SAs this way?

Keeping with the same example we have been using, if an SA has only reached 70%, there is a pretty good chance that we will find 13 or more of 19 women in one sample who know two ways that HIV spreads and, thus, conclude that it has reached the target when in fact it has not. In fact, with a single survey we would misclassify this SA wrongly 67% of the time.[41] Again, in multiple surveys of the same SA (if there were no change) the chance of misclassifying the SA twice goes way down. If an SA has only reached 60%, we will still find 13 out of 19 in a single sample about one-third of the time and wrongly classify the SA as having reached the target when it clearly did not. It is only if the SA has only reached 50% that we become highly unlikely to find 13 of 19 and classify it (wrongly) as having reached the target. In fact, if the true SA percentage is 50% (the SA is a "50 percent community"), we would misclassify it in a single survey less than 10% of the time—a level of error we are more comfortable with.[42]

We come back to the same point we have made before: LQAS allows us to accurately identify SAs that are falling far below the target with a fair deal of accuracy. For those in between we

[40] This is called beta error. See the section below "LQAS Applications: Special Topics" for more details on this.
[41] Again this statement is based on the cumulative binomial distribution of probabilities. In this case we know that the probability of finding 13 OR MORE who know two ways to prevent the spread of HIV in a sample of 19—from a population in which the true percent who know two ways is **70%**--is 67%.
[42] This is a key to understanding how LQAS has been set up—and that we develop more below: the probability of misclassifying a successful SA as unsuccessful is <10% AND, the probability of misclassifying an SA that falls 30 percentage points below the target as "successful" is also <10%. The "gray area" in between is of less interest to us if we can accurately identify those who are falling far below the target.

may classify them wrongly. They will be classified as having met the target, while they are short of it. But managers may be willing to live with that because they need to focus on those SAs that are in greatest need of support (and most of the time—90% of the time—those who fall 30 percentage points below the target will not make the cutoff).

IS IT WORTH THE TROUBLE?

Given all the possible problems of misclassification, you may wonder whether it is worth using LQAS. It is important to keep a few things in mind before deciding:

- As we have said, if we are using LQAS in a routine way, the chance of misclassifying an SA multiple times is very small.
- There is very little chance of misclassifying (and discouraging) successful SAs using LQAS.
- We do not spend—and potentially waste—resources precisely identifying SAs that are below the target but not too far below.
- When we select SAs that are below the decision rule for remedial interventions, we are likely to have included those that are falling far below the target.

Before moving on into greater detail, the following table shows the decision rules for target percentages from 10% to 95% in increments of 5 percentage points. Note that the decision rule for 80%, as we have seen, is 13, based on a sample of 19. Its use is simple. For example, imagine that I am responsible for a program zone (an SA) and had another indicator such as the percentage of caregivers who gave their child more to drink when they had diarrhea. Suppose in this case the program target is that 60% of caregivers should be practicing this by the end of the program. If I find that at least 9 caregivers out of the sample of 19 in my SA give their children more to drink when they have diarrhea (obtained by asking the caregivers what they did when their child had diarrhea in the past 2 weeks), I classify my SA as having reached the target. If my target is 40%, then I would need at least 5; 85% at least 14; and so on. We present a more complete table of decision rules for sample sizes other than 19 below.

Table M3.1: Decision rules for a sample size of 19 at each target level for coverage

	Coverage target* for indicator (percentage) in SA																	
	10%	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**	**45%**	**50%**	**55%**	**60%**	**65%**	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**
Decision rules for n=19	--	--	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Logic statement	"In one SA, if the true population coverage level is above the desired target (i.e. 80%), then the number of respondents out of 19 who respond correctly will be equal or greater to the decision rule figure (i.e., 13) 93% of the time. With this decision rule *we will classify the SA as having reached the target of 80%* (even though we know that it may have fallen below the target somewhat). What we are fairly sure of is that we have not wrongly classified an SA at 50% (or below) of having reached 80%"																	

* Options for a baseline survey (no pre-identified target) are discussed below.

RESPONDENT SELECTION PROCEDURES IN LQAS—A PRECISION

We have repeated several times that the selection of our sample in LQAS uses the same procedures as already outlined for cluster sampling. This means that we can use the same procedures for selecting communities—from a list, with probability of selection proportionate to the size of the community. We can also use the same procedures for selecting households and respondents. However, there is one important difference between LQAS and cluster samples. In LQAS, we must develop a community list and conduct the community selection process for each SA. In cluster sampling, we place all communities in a program area in one list and select 30 clusters from that list. In LQAS, we place all communities *in each SA* in a separate list and use the same procedure and select 19 interview sites in each SA. Thus, instead of selecting 30 total clusters and interviewing 10 individuals per cluster (as we discussed above), we select 19 interview sites in each SA. Further, generally, we only interview one respondent in each of the 19 interview sites in LQAS.

However, recall that in cluster sampling it is possible to select multiple clusters in a single community. In a similar way in LQAS, it is possible that we will select multiple interview sites in a single community. For example, if we have only 15 distinct communities in an SA, then it is clear that in some of the larger communities we will select more than one interview site, implying that we will interview more than one person in that particular community. If that is the case, we go back and repeat the mapping process we used to select the first respondent in that community. In such a case we recommend that you exclude the section in which you interviewed the first respondent and randomly select one of the other sections and map it in order to select a household. In other words, we start the mapping process over as if we had just entered the community while excluding the section we previously selected. We repeat this process for each individual interview we conduct in an interview site.

The terminology distinguishing 30 cluster and LQAS approaches can be confusing. Table M3.2 seeks to sort out the differences in terminology used in each.

Table M3.2: Key differences in cluster and LQAS sampling strategies

	30x10-cluster sample	LQAS
Where we begin the sampling procedure	A list of all communities (with their population) for the entire program area	A list of all communities (with their population) for each SA
What we select in the first stage of sampling	30 clusters (in some large communities we may select more than one cluster)	19 interview sites in the SA (in some large communities we may select more than one interview site)
What we select in the second stage of sampling	10 Households with eligible respondents per cluster	One household* with eligible respondent per interview site
What we select last	One respondent in each of 10 households selected in each cluster (based on what indicators call for)	One respondent in the household selected at the interview site (based on what indicators call for)
The result	10 completed interviews for each cluster	1 completed interview for each interview site*

*There are some exceptions to this as we have seen and will see in more detail for LQAS below.

SPECIFIC APPLICATIONS OF LQAS

LQAS APPLICATIONS: GENERAL MONITORING

In this section we return to an LQAS application example and use it to reiterate what a sample of 19 is and is not good for in greater depth. The following map shows an entire program area and the area is divided into five SAs. We are conducting a given intervention in an entire program area, but we want to assess its progress in SAs.

An SA should be a meaningful program area and not merely an artificial grouping. Ideally, it represents a unique geographical area administered by a team exclusively responsible for implementation in that area. However, an SA can also be defined for geographical, language, or ethnic reasons, or any variable that might lead to differing results.

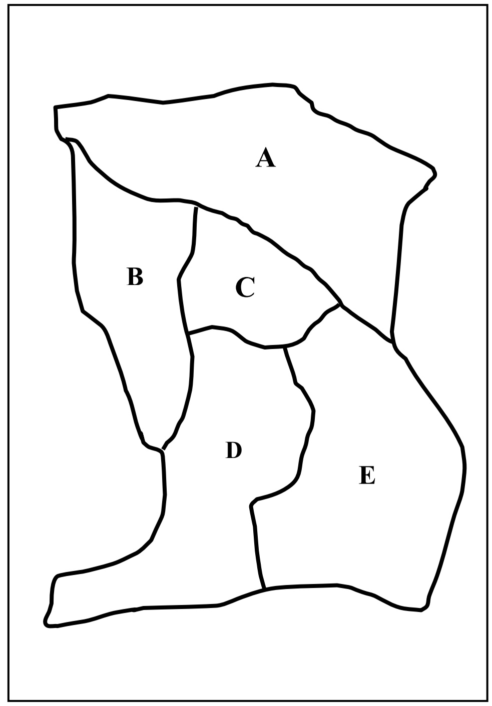


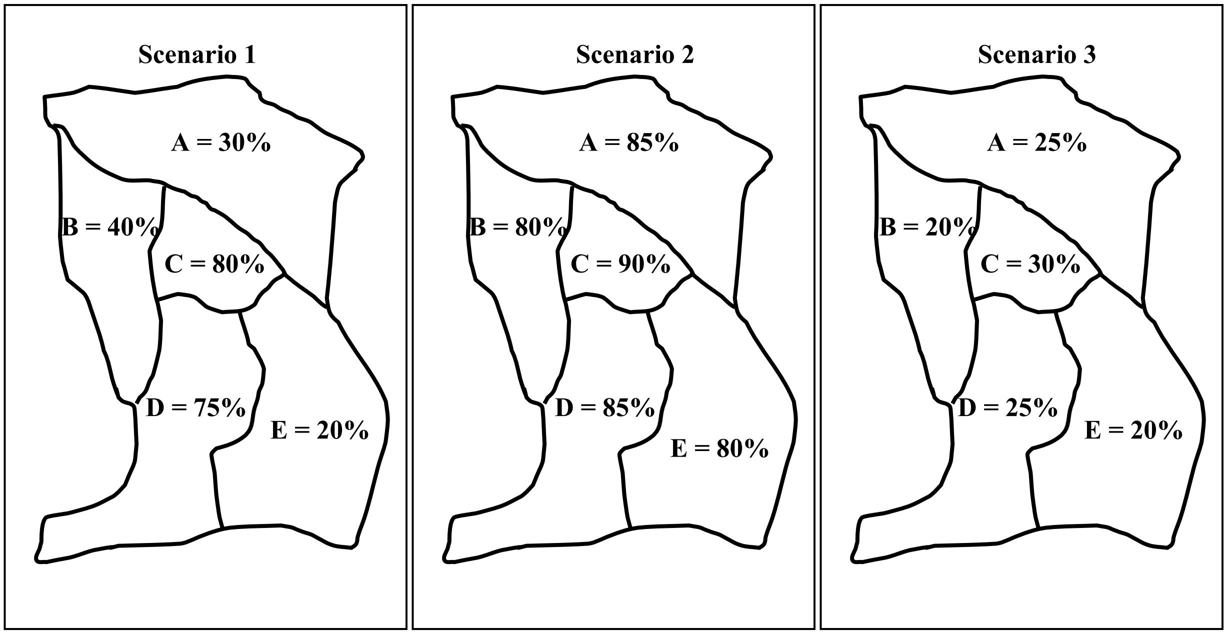
Figure M3.2: Program area and supervision areas

Together, A, B, C, D, and E represent the Program Area. A, B, C, D, and E represent five SAs.

Let us assume that five different sub-teams are each responsible for a single supervision area: A, B, C, D, or E (see Figure M3.2). LQAS is designed to give each team some information about how well it is doing in reaching targets. It also helps overall program managers ascertain whether certain SAs are falling far behind or whether certain SAs, because of their success, might be resources for other zones. Let's assume that our target for this indicator is 70% (decision rule of 11 as per Table M3.1—please verify this for yourself). We will now consider three scenarios for this program area. The percentages given represent the true population values for the indicator: percentage of women who know two ways to prevent the spread of HIV. Obviously, we never know this true proportion in the real world (or we wouldn't need to carry out surveys), but the mental exercise of imagining that we do will help us demonstrate when LQAS might be most useful.

In Scenario 1 of Figure M3.3, we see that two of the SAs—C and D—are performing much better (and reaching the target), in relation to this indicator, than SAs A, B and E. We know that C and D would be very likely to have 11 or more women knowing two ways to prevent the spread of HIV, and that A, B, and E would be unlikely to reach the same threshold. Consequently, a survey based on LQAS would enable us to identify the difference between the two types of SAs and to strategize about learning from C and D to better support A, B, and E.

Figure M3.3: True coverage estimates in five supervision areas according to three scenarios



In Scenario 2, we see that all SAs are reaching the target. We have already seen that there is some chance of misclassifying SAs but, in this case, given that each of the SAs is far surpassing the target, we would likely conclude, based on a sample of 19, that each of the SAs is reaching the target. In this case identifying differences between the SAs will not be possible. This is not a problem since all SAs are reaching the target.

In Scenario 3, we see that none of the SAs is reaching the target. Again, given that they are far below the target, there is virtually no chance that we will misclassify them has having reached the target with a sample of 19. In this case, using LQAS would lead us to conclude that none of the SAs is reaching the target, but it would, again, not enable us to identify differences between them.

You can imagine other scenarios in which we would be more likely to misclassify SAs in one way or another with LQAS. What these three scenarios illustrate is that LQAS is MOST useful in cases in which there are large differences between SAs—with some doing well and others falling far below. This is not to say that LQAS is not useful in the other cases. As we shall see below, we can combine the results of all five LQA samples to estimate, fairly precisely, the indicator for the entire program area. This is the case in ALL scenarios. In addition, this example shows only one indicator. Typically there would be several indicators of interest, and differences among SAs may vary with the selected indicator.

Table M3.3 summarizes this set of examples concerning the use of LQAS.

Table M3.3: Information produced by LQAS in different situations

When there are large differences in coverage among supervision areas, LQAS enables us to:	Identify the low-coverage SAs to be able to:	• Learn causes of low coverage • Focus our efforts and resources on these SAs • Improve coverage of the whole program area by improving coverage in these SAs
	Identify high-coverage SAs to be able to:	• Study and learn what is working well • Identify things that can be applied to other SAs
When there is little difference in coverage among SAs, the use of LQAS can still help us in the following ways:	If coverage is generally high,	• It may help us to think about shifting resources to improve other areas of knowledge and practices.
	If coverage is generally low it enables us to think about what we can do to:	• Study causes of low coverage • Identify/study other knowledge or practice areas that ARE reaching targets to learn what is working well

LQAS Applications: Estimating Coverage for an Entire Program Area

LQAS can be used both to help SAs determine how well they are doing in relation to targets set for coverage, but also for estimating overall program average coverage. We have seen generally how the first use works and how we can use information at both the SA and program level, since there is little reason to use LQAS if one is not interested in this management level of information. Here we deal briefly with how to combine LQA samples from all SAs to estimate coverage for an entire program area. In the next section, we will discuss how to use this information at baseline.

Let us continue with the same indicator we used previously: the percentage of women who know two ways to prevent the spread of HIV. Once we have conducted our surveys in all SAs, we can combine the results to estimate coverage for the entire program area.

The following table shows an example, which we will assume comes from a final evaluation whereby we wanted 80% of women to know at least two ways to prevent the spread of HIV by the end of our project. Here are the results of surveys in the five SAs we saw previously for this indicator (Table M3.4). Keep in mind that we would typically have more than a single indicator.

Table M3.4: Final evaluation LQAS respondent results for five SAs

SA	Number of respondents out of 19 who know two or more ways to prevent the spread of HIV	Sample size	Coverage estimate
A	15	19	Not appropriate
B	11	19	Not appropriate
C	16	19	Not appropriate
D	13	19	Not appropriate
E	17	19	Not appropriate
Total	**72**	**95**	**75.8%**

Our total sample size across all SAs is 95. To get the average coverage for the entire program area, we divide 72/95 which equals 75.8%.[43] We can also use the LQAS table for sample size 19 (see above) to assess whether individual SAs have reached the target of 80%. Assess for yourself whether all SAs have done so (hint: find the decision rule for 80% above first).

Given the above result, what might we conclude about whether our program has reached the target? Since we have taken a sample (and have used random sampling procedures), we can and should calculate a confidence interval around our "point estimate." Such intervals (as we have seen above in the cluster sampling module) allow us to say how certain we are that our results represent the true percentage in the population by giving a range in which we assume the true population result lies.

As a reminder, the formula for calculating a 95% confidence interval is:

- Confidence interval = $p \pm 1.96 * \sqrt{((p * q)/n)}$
  Where p=the percentage from our sample expressed as a simple proportion,
  1.96 comes from a standard statistical table relating to 95% confidence,
  q=1-p and
  n=our total sample size.

In our example we would have:

- Confidence interval=$.758 \pm 1.96 * \sqrt{((.758 * .242)/95)}$
- Our 95% confidence interval is thus 67.2%-84.4%.

What this means is that we are 95% confident that the true population percentage for this indicator lies between 67.2% and 84.4% for our program area.

[43] In the final part of the LQAS section we will revisit this procedure and provide some more details on some modifications we may need to make to ensure the accuracy of our estimate of overall program coverage and our confidence interval. The modifications are necessary because the total populations in different SAs could be quite different (some SAs will be relatively small and others relatively large), and we must "weight" the results to account for the fact that some are larger and should, therefore, have a greater influence on the overall percentage estimate (while others are smaller and should have a smaller influence on the overall percentage).

Given this, even though our POINT estimate from the survey of all SAs is 75.8%—below the program target of 80%—we cannot conclude that we have not reached the target based on this sample.

We will see later how to improve our overall estimate by taking into consideration the size of each SA (referred to as "weighting" the results), but for now, the foregoing has shown us how we can use the results of individual LQAS samples in each SA to calculate the average coverage for the entire program area in SAs of similar sizes.

LQAS APPLICATIONS: BASELINE SURVEYS

To this point we have focused on the use of LQAS for monitoring and evaluation purposes, that is, 1) when we have a set target for an indicator and we want to assess whether we have evidence that SAs have reached the target and 2) to estimate coverage for an entire program area based on the samples of each SA to see how we are doing program wide in relation to that indicator. In addition to these uses that we employ after a program is underway or at program's end, we can use LQAS at baseline not only to assess coverage for the entire program area (which we just described), but also to see whether there is any evidence—even before the program begins—that certain SAs have much lower levels of coverage—compared with other SAs—in relation to the indicators of interest to our program. Again, we return to an example to illustrate this.

Suppose that we again take 5 samples of 19 respondents—before our program begins—in each of our SAs, and find the following results (see Table M3.5) in relation to the indicator: percentage of women 15-49 who know two ways to prevent the spread of HIV.

How might we use these results to draw some initial conclusions about important differences between these SAs in relation to this indicator? Recall that it is NOT appropriate to calculate individual percentages for each SA because of the small sample size in each one.

We *can*, however, use an estimate of coverage for the overall program area, find the decision rule that corresponds with that estimate, and use it to see whether a given SA seems to be far below that level at baseline. The procedure works as follows:

1. First, create an estimate of coverage for the entire program area by summing the number of respondents who know two ways to prevent HIV transmission and dividing it by the total sample size.

   As seen below, our program area estimate is 65.3% (i.e., (62/95)*100).

Table M3.5: Baseline survey result in a program area of five supervision areas

SA	Number of respondents out of 19 who know two or more ways to prevent the spread of HIV	Sample size	Coverage estimate
A	12	19	Not appropriate
B	9	19	Not appropriate
C	16	19	Not appropriate
D	11	19	Not appropriate
E	14	19	Not appropriate
Program area	**62**	**95**	**65.3%**

2. Next find the decision rule for this percentage for this sample size. We round up our program area estimate of 65.3%[44] to nearest 5% (70% in this case). Using the table below we see that for a sample size of 19 the decision rule for 70% is 11.

Coverage target or average coverage (for baselines)																		
Percentage	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
Decision Rules for n=19	--	--	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

3. We can now ask: Are there any SAs at baseline whose coverage appears not to be above the average for the entire program area? This means identifying the SAs that fail to be at or above the decision rule.[45]

Table M3.6 provides the results for the current example, answering the question, "Is an SA at or above the average coverage?" Based on the answer, as seen in Table M3.3, we will make different management decisions for each SA.

[44] Again, we may need to "weight" this result. We describe how to do this below.
[45] We know from our introduction, that for SAs that are close to the program area average, it is possible that a number may be coming up above the decision rule, but we also know that this is very unlikely to happen for SAs that are far below the program average.

Table M3.6: Establishing supervision areas with higher needs based on a baseline survey

SA	Number of respondents out of 19 who know two or more ways to prevent the spread of HIV	Sample size	Coverage estimate	Decision rule for classifying SA as having reached the program coverage average (rounded up to 70%)	Equal to or above average coverage?
A	12	19	N/A	11	Yes
B	9	19	N/A	11	No
C	16	19	N/A	11	Yes
D	11	19	N/A	11	Yes
E	14	19	N/A	11	Yes
Program area	62	95	65.3%		

4. This analysis allows us to make a management decision. We conclude that SA B does not have coverage for this indicator that is equal to or above the average for the entire program area. This signals to us that B may merit some further analysis and/or attention in order to understand why levels of knowledge for this indicator at baseline are below the average.

We have just considered how LQAS can be used at baseline as a signal to alert us to important differences among SAs (should they exist) before the program begins. The following table takes a different tack—still based on the same baseline survey—to show how a single SA (B in this case) might use information about a variety of indicators measured at baseline. The program still concerns reducing the transmission of HIV and concerns both men and women.

We report the same indicator as in the previous example, but we also asked men certain questions about their knowledge and practices and asked women other questions as well. We calculated the program coverage estimates for each indicator in the same way as in the example above and found our decision rules from the same table for sample size 19 as we saw previously (be sure to consult it to confirm that we chose the correct decision rules).

Table M3.7: Reviewing performance on multiple indicators in one supervision area

Indicator SA B	Correct out of 19	Program coverage estimate	Decision rule (from sample size of 19)	Equal or above average coverage?
		→		
Women 15–49 who use condoms when having sex	7	45%	**6**	Yes
Men 15–49 who use condoms when having sex	4	20%	**1**	Yes
Women know two ways to prevent HIV transmission	9	70%	**11**	No
Men know two ways to prevent HIV transmission	13	65%	**10**	Yes
Women who know where to get an HIV test	6	30%	**3**	Yes

Since this SA has results equal to or above the average coverage for all other indicators we may want to reflect on what, if anything is particular to this indicator or whether this is merely an example of a sampling error. Whatever the case, this example shows how a program and SA can use information generated from baseline surveys to examine SA characteristics vis-à-vis program indicators even before the program begins.

LQAS Applications: Putting It All Together for Monitoring and Evaluation

In this section, we put the all the foregoing topics together to demonstrate how to move from questionnaires, to questionnaire tabulation in an SA, to summary tabulations, to using LQAS over the life of a project to assess indicators over the life of a project. We will simplify things by using the same indictor as in the above examples.

Let us begin with the questionnaire. Assume that we have already used our sample selection process and have identified women age 15–49 as respondents. Table M3.8 contains questions from an actual survey that we could use to assess the indicator.

Table M3.8: Interview survey questions

NO.	QUESTIONS AND FILTERS	CODING CATEGORIES		SKIP
1.	Now I would like to talk about something else. Have you ever heard of an illness called AIDS?	YES ………………………………. NO ………………………………..	1 2	→ Skip to the end
2.	Can people reduce their chances of getting the AIDS virus by having just one sex partner who is not infected and who has no other partners?	YES ………………………………. NO ……………………………….. DON'T KNOW …………………..	1 2 88	
3.	Can people get the AIDS virus from mosquito bites?	YES ………………………………. NO ……………………………….. DON'T KNOW …………………..	1 2 88	
4.	Can people reduce their chances of getting the AIDS virus by using a condom every time they have sex?	YES ………………………………. NO ……………………………….. DON'T KNOW …………………..	1 2 88	
5.	Can people get the AIDS virus by sharing food with a person who has AIDS?	YES ………………………………. NO ……………………………….. DON'T KNOW …………………..	1 2 88	
6.	Can people reduce their chance of getting the AIDS virus by abstaining from sexual intercourse?	YES ………………………………. NO ……………………………….. DON'T KNOW …………………..	1 2 88	

NO.	QUESTIONS AND FILTERS	CODING CATEGORIES		SKIP
7.	Can people get the AIDS virus because of witchcraft or other supernatural means?	YES ……………………………… NO ………………………………. DON'T KNOW …………………..	1 2 88	

In this example (these questions are only part of a larger set), we could decide that if the respondent answered "Yes" to at least two of the three questions about sexual relations (Questions 2, 4, and 6) we would consider that he or she knows at least two ways to prevent the spread of HIV. This could be our way of operationalizing the indicator. Another program might decide that, not only must the respondent answer "Yes" to two of these three questions, but he or she must also answer "No" to Questions 3, 5, and 7. This is a slightly different indicator because it is really "knows at least two ways and correctly identifies erroneous information." The point is that the decision about how to operationalize the indicator can vary.[46]

In our case, we will use the first choice—that the respondents answer "Yes" to at least two of the three Questions 2, 4, and 6. Keep in mind that if someone responds "No" to Question 1 ("Have you ever heard of an illness called AIDS?"), we will not ask them any of the other questions (notice the questionnaire's directive: "Skip to the end"). In this case, we could simply decide that if the person answers "No" to Question 1, then he or she does not know at least two ways to prevent the spread of HIV.

Once we have carried out our survey with 19 women age 15–49 in our SA, we can collect all the questionnaires and tabulate the results for this indicator in order to analyze it for our SA and combine our results with those from other SAs to estimate coverage for this indicator for the entire program area. Our tabulation table might look like Table M3.9, and we could produce it by computer or by hand.

Again, this is an abridged table, and normally we would have many indicators and the table could run on for several worksheets. Notice that we have included coding for question 1 in the table because it feeds into the indicator of interest. We may not analyze this separately, but it helps us with quality control, meaning that each time a respondent says that he or she has never heard of HIV (0 for Indicator 1), he or she should also have a 0 for Indicator 2 in the table—verify that this is the case.

Notice also that the second column of the table provides us with a brief summary of how to decide whether to put a "0" or a "1" in the column and corresponds to the decision we named above. We have filled out the table using the example from SA B above. The key is that we take our results directly from the questionnaires (such as the example in Table M3.8).

Examine the completed Table M3.9. You will notice a few things:

- We do not have the questionnaires here, but the information was transferred from them to this table following the rules laid out in the "Correct Response Key" column.

[46] Our point is not to decry the standardization of indicators the many times when it is appropriate, but to let the reader follow clearly the production of an indicator from a set of interview questions.

- While 11 respondents in this SA (out of 19) had heard of HIV/AIDS, only 9 of them correctly identified two ways to prevent its transmission.
- In each case in which the respondents had not heard of HIV (or said that they did not know), they received a zero for the first indicator AND for the indicator we are interested in. Please verify this. In other words, if a respondent answers “No” or “Do Not Know” for the first question, he or she could not get a “1” for the second indicator.
- There were no missing responses, so our total sample size was 19 (we discuss below options of what you can do if there ARE missing responses).

Table M3.9: Result tabulation table for a supervision area

Baseline Survey and Regular Monitoring: Females 15–49 Years																							
Supervision Area: ________B____________									Supervisor: ______________________										Date: _____________				
CORRECT = 1							INCORRECT = 0								MISSING = X								
			Interview/Questionnaire Number																				
#	Indicator	Correct Response Key	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total Correct in SA	Total Sample Size (all '0's and '1's)
Section X: HIV/AIDS																							
1	Have you ever heard of an illness called HIV/AIDS?	"Yes"=1 (if "No" or "Unknown" then indicator Number 2 is automatically incorrect=0)	1	0	1	1	0	0	1	1	0	0	0	0	1	1	1	0	1	1	1	11	19
2	What a person can do to prevent the spread of HIV/AIDS	2 or more "Yes" to Questions 2, 4 and 6=1	1	0	1	1	0	0	1	1	0	0	0	0	1	0	1	0	1	0	1	9	19

We have now tabulated responses for our SA for this indicator. We had already seen the result for SA B previously, and these results correspond to what we have seen. Notice that the table title indicates baseline survey or regular monitoring. The same table could be used for a final evaluation as well. Now that we have completed the work, the next question is, what do we do with the information?

The answer to that question depends.

- If this is a baseline survey, then we will want to combine these results with the results of other SAs in order to see if this SA has fallen below the average coverage for this indicator for the entire program area (see prior example).
- If this survey is done for routine monitoring, we still may want to combine these results with others and assess where this SA is in relation to the other SAs and in relation to the target we have set for each SA.

Let us consider this routine monitoring use in greater detail.

Suppose that at baseline (this differs a bit from the earlier example) we found that the average coverage for the entire program was 50%—that is, 50% of women age 15–49 interviewed in all SAs at baseline were able to name at least two ways to prevent the transmission of HIV. As a program, we want to bring that level of knowledge up to 85% by the end of our program (which runs for four years). We have also set some interim targets, and by the end of year 3, our target is that, program wide and in each SA, 75% of women age 15–49 will know at least two ways to prevent transmission of HIV. We are now at the end of year 3 and have just conducted our survey in each SA. We have already seen the results for SA B. We can use Table M3.10 to summarize the results for all SAs for this indicator. This table is different than Table M3.9, and we will describe how to use it on the following pages.

Table M3.10: Summary tabulation table: Example 1

Baseline Survey: Females 15–49 Years

Program name: ____________________ **DATE: _______________**

#	Indicator	Total Correct in Each SA and Decision Rules for Average Coverage and Target Coverage					Total Correct in Program	Sample Size					Total Sample Size in Program	Average Coverage = Total Correct Sample Size	Target Cover-age
		A	B	C	D	E		A	B	C	D	E			
Section X: HIV/AIDS															
1	Have you ever heard of an illness called HIV/AIDS?														
2	What a person can do to prevent the spread of HIV/AIDS														

Before filling it in, let us look at the features of this table. Unlike the previous tabulation table, this one has room for summary information for all five SAs. Thus, information for this table comes from tables like Table M3.9 for each SA (we only showed the one for SA B in M3.9). As in that case, we only have one indicator shown here. Typically this table would contain all key indicators for the program. Let us focus on the same indicator as in the previous table. Here, however, we have two lines for each indicator, and the second line is split into two columns for each SA. On the first line we will place the total correct for each SA from the previous table. We already know the total correct for SA B—it is 9.

We are going to add one level of analysis for this example. The second line in table M3.10 has two cells. These cells are used to record two different decision rules that we can take from the LQAS table.

- The first cell will be for the decision rule for the program target that we have set. As noted above, this is the third year of the program, and our target coverage for this indicator in year 3 is 75%. Referring back to the LQAS table for sample size 19 (which is the sample obtained for this variable in each SA), we note that the decision rule for 75% is 12. We would place that in the left-hand lower cell for each SA for which we had a sample of 19 (all of them in this example).
- In the second cell on the bottom row, we will place the decision rule corresponding to the average coverage for the entire program area based on all SAs. Even though we are not at baseline but in year three, the approach is the same as described above.

Why would we have two decision rules: one for the target coverage (75%) and one for the average coverage of the entire program? We do this because this allows us to analyze our SA in two ways:

- We can ask: “Can we classify our SA as having reached the target?” (Yes or No)
- We can ask: “Can we classify our SA as having reached the average coverage of all the SAs combined?” (Yes or No)

Let us fill in the table and see what we can learn from an example. We will simply ignore the first indicator at this point and focus on the indicator of interest to us.

Table M3.11: Summary tabulation table: Example 2

<table>
<tr><th colspan="22">Baseline Survey: Females 15–49 Years</th></tr>
<tr><th colspan="16">Program name: ____________________</th><th colspan="6">DATE: _______________</th></tr>
<tr><th rowspan="2">#</th><th rowspan="2">Indicator</th><th colspan="10">Total Correct in Each SA and Decision Rules for Target Coverage [Left Cell] and Average Coverage* [Right Cell]
* rounded up to near 5%</th><th rowspan="2">Total Correct in Program</th><th colspan="5">Sample Size</th><th rowspan="2">Total Sample Size in Program</th><th rowspan="2">Average Coverage = Total Correct / Sample Size</th><th rowspan="2">Target Cover-age</th></tr>
<tr><th colspan="2">A</th><th colspan="2">B</th><th colspan="2">C</th><th colspan="2">D</th><th colspan="2">E</th><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th></tr>
<tr><td colspan="22">Section X: HIV/AIDS</td></tr>
<tr><td rowspan="2">1</td><td rowspan="2">Have you ever heard of an illness called HIV/AIDS?</td><td colspan="2"></td><td colspan="2"></td><td colspan="2"></td><td colspan="2"></td><td colspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">What a person can do to prevent the spread of HIV/AIDS</td><td colspan="2">12</td><td colspan="2">9</td><td colspan="2">16</td><td colspan="2">11</td><td colspan="2">14</td><td rowspan="2">62</td><td rowspan="2">19</td><td rowspan="2">19</td><td rowspan="2">19</td><td rowspan="2">19</td><td rowspan="2">19</td><td rowspan="2">95</td><td rowspan="2">65.3%</td><td rowspan="2">75%</td></tr>
<tr><td>12</td><td>11</td><td>12</td><td>11</td><td>12</td><td>11</td><td>12</td><td>11</td><td>12</td><td>11</td></tr>
</table>

Let us examine these results and note the following key points:

- Before going on to SA-level analyses, let us note that, overall, the entire program may not appear to be meeting the target coverage that we have set. Now it is important to note that this conclusion is based on the point estimate of 65.3% for the entire program area. However, if we calculate a 95% confidence interval for this point estimate, with a sample size of 95, the point estimate would be in the range of 55-75%. Consequently we are not able to conclude (based on this confidence interval) that we are not reaching the target of 75%.
- SAs A, C, and E have reached both the decision rule for both the target and for the average coverage. This means that, at the end of the third year, we classify these three SAs as having achieved the target of 75% coverage for this indicator (as we have seen, what we're really sure of is that these SAs are not below 45%). We can also conclude that they are doing as well as or better than the overall program coverage.
- SA D has not reached the target coverage decision rule, and we conclude that SA D has not, therefore, at the end of three years, reached the target that we set for it and the entire program. However, it has reached the decision rule for the overall program coverage area, and therefore we can conclude that it is doing no worse than the overall program.
- Finally, we see that SA B has not reached the decision rule for either the target or the overall program coverage level. Thus, we conclude that SA B is not only not reaching the target that we have set for it and the entire program, but it is also falling below the overall coverage average for the entire program area. Clearly, we need to find out what is happening in SA B.

While we might be somewhat concerned about SA D, we should probably focus first on SA B. This conclusion leads us to a final summary table that describes how we might use the results of LQAS routine monitoring to prioritize interventions in SAs that are falling short of targets or overall average program coverage for an indicator. This table helps us map out where our five SAs fall.

Table M3.12: How to identify priority supervision areas using the summary tables during regular monitoring

SA classification result on coverage target	SA classification result on average coverage for program area	Priority for corrective action [supervision area from our example]
Below	Below	Highest [B]
Below	Not Below	Second Highest [D]
Not Below	Below	Second Highest
Not Below	Not Below	Not a Priority [A, C, E]

In this section, we have not explained in detail all the procedures necessary for survey design, interviewing, developing indicators, tabulating all results, developing an analysis protocol, or summarizing results. These points go beyond the scope of this manual, but a link to information about an effective training manual, with handouts and lesson plans already designed, is provided in Annex 1: Key Resources. We strongly recommend that you use these training materials to train staff conducting LQAS surveys before they conduct the surveys. Many of the examples in this section are taken from this training manual.

A note of caution: It is also important to restate a fundamental issue: all this is based on fully respecting strict principles for random selection of respondents. In some cases—for example, when repeat LQAS surveys are carried out by project staff in their SA—interviewers can be tempted to avoid having to drive to a new place to interview the 17th, 18th, or 19th respondent. As non-specialists, they may feel that they could just "get another random respondent" in this or that village. That kind of practice, however, will totally invalidate—i.e., lead to misleading findings—the process on which LQAS resides. Given the small size of the samples, what may seem like marginal liberties with the process and its principles can have a very negative impact on the validity of the findings.

One final note: Throughout this module we have focused on comparing LQAS results for an entire program area (combining the results from all SAs) to a pre-defined set of targets for key indicators. As we discussed in the two-stage cluster survey section if you want to compare coverage estimates at two points in time (baseline and end of program) you will need to adjust your sample sizes accordingly. If you do want to do this using LQAS the implication is that you may need to increase sample sizes within SAs above 19. We discuss the issue of using LQAS with sample sizes other than 19 below but keep in mind that increasing sample sizes within SAs (in order to increase the precision of program-wide estimates) will increase the costs of the survey.

We now move onto some special issues encountered in surveys and how to deal with them when LQAS is the chosen sampling methodology.

LQAS APPLICATIONS: SPECIAL TOPICS

WEIGHTING RESULTS WHEN ESTIMATING COVERAGE FOR AN ENTIRE PROGRAM AREA

At the end of the section, LQAS Applications: Estimating Coverage for an Entire Program Area, we alluded to the fact that when estimating coverage for an entire program area it might be necessary to adjust the estimate (from, say five SAs) to account for differing population sizes in different SAs. Here we explain this in more detail and provide examples to demonstrate the potential importance of these adjustments.

The process of adjustment is referred to as weighting because we want to make sure that the influence of each SA's results on the average coverage estimate is commensurate with its relative size, with larger SAs having more influence on the overall estimate—as they should if we are saying that the estimate represents the entire population of the program area—and, conversely, smaller SA's having less influence on the overall estimate. Weighting is not needed when making LQAS judgments of a single SA.

Data can be weighted by SA population sizes using the direct adjustment method. We will illustrate this method here. Let's assume that 19 interview sets were carried out in each SA, regardless of the SA population size and the number of SAs in the program area. Without weighting, a sample of 19 can potentially either overestimate or underestimate the coverage estimate. Weighting the data allows us to remove this distortion.

Let's look at Table M3.13, which provides the details on how to weight results for an indicator. We will stick to the familiar example of "percentage of women age 15–49 who know at least two ways to prevent the transmission of HIV" and return to the five SAs we have been looking at throughout this section. The correct responses are different from those above but let us not be concerned about that—we can assume that this is a completely different sample at another time.

Table M3.13: Weighting the results of an indicator

Percentage of women 15-49 who know at least two ways to prevent the transmission of HIV						
1	2	3	4	5	6	7
Supervision Area	Sample Size n	Correct Responses c	Mini-Proportion p=c/n	Population Size N	Weight Wt=N/∑N	Weight*Mini Percent Wt*p
A	19	11	0.579	10,718	0.245	0.142
B	19	12	0.632	6,379	0.146	0.092
C	19	8	0.421	9,379	0.215	0.090
D	19	4	0.211	9,731	0.223	0.047
E	19	7	0.368	7,500	0.172	0.063
Totals	95	42	0.442	43,707		0.435

- Column 1 lists the five SAs in the program area.
- Column 2 is the sample size of each SA (19 in this case), with the total sample size of 95 at the bottom in the "Totals" row. The symbol for mathematical equations that use the results of this column is "n."
- Column 3 is the number of correct responses in this SA—that is, the number of women in the sample of 19 who correctly gave two ways to prevent the transmission of HIV. The symbol for mathematical equations that use the results of this column is "c."
- Column 4 is what we will call the "mini-proportion" for each SA. The symbol for mathematical equations that use the results of this column is "p," and we can see that p is equal to the number of correct answers in each SA divided by the sample size for that SA (p=c/n). Please note that we never report individual SA proportions and do so here only to calculate a weighted average for the entire program area.
- Column 5, which would come from census or other data (and was probably used to select interview sites when you developed your sample), is merely the total population of each SA. It need not be the population of women age 15–49 because we assume the proportion of women 15-49 is the same in each SA. This assumption is reasonable. Notice that SA A is the largest SA in terms of population (and, thus, women age15–49) and that SA B is the smallest. The symbol for mathematical equations that use the results of this column is "N" and the symbol ∑N means the sum of all the individual SA "N." Thus ∑N=43,707—the final number in Column 5.

- Column 6 is critical to our understanding of how we calculate weighted proportions because it is the weight that will be applied to the mini-proportions to allow for larger SAs to influence the adjusted coverage average more (and smaller SAs to influence it less). The symbol for mathematical equations that use the results of this column is "Wt," and we can see that Wt is the N for a given SA divided by the ∑N (the sum of all the individual SA N). Thus for SA A, Wt=10,718/43,707=.245. Please verify that the relative amount of the weight for each SA corresponds exactly to its population size, with the largest SA having the greatest Weight.
- Finally, Column 7 takes the Weight for each SA and multiplies it by the mini-proportion (Column 6 * Column 4). Thus, this column is Wt*p. So, for SA A, Column 7 is .245*.579. The individual rows of this column hold no interest to us—only the Totals row matters where we add everything up.

Let us turn to the Totals row now to examine the results of our calculations. If we look at the Totals row for Column 4, we have the unweighted estimate for the entire program area for this indicator. The total for this column is calculated using the approach we described in the section, LQAS Applications: Estimating Coverage for an Entire Program Area. It is given as a proportion here, but if we multiply it by 100 we get the percent average coverage for the entire program area: 44.2%. This means that our point estimate of the average coverage for this indicator in our program is 44.2%—we estimate that 44.2% of women age 15–49 in our program areas know two ways to prevent HIV transmission. We discussed how to calculate a 95% confidence interval for this number but will not do that here. We will calculate a confidence interval for our weighted coverage below.

Notice that the proportion in the Totals row for Column 7 is .435 (43.5%). You may say to yourself (correctly in this case) that weighting has not changed the percentage very much. That is because the differences in the sizes of the SAs are not (relatively speaking) very large. However, we can ascertain that the weighted average is a bit lower due to that fact that some of the lower performing SAs (D notably) had relatively larger populations and thus "pulled down" the average when given more weight (as they should) than SA B, for example.

Table M3.14 provides a much clearer example of how different population sizes can affect the weighted average. To make this clear, we keep everything constant from the first example (total population N, correct responses by SA, and sample sizes). What we do alter is the population size of each SA to demonstrate how population can affect results.

Table M3.14: How population sizes can affect the weighting of results of an indicator

Percentage of women age 15–49 who know at least two ways to prevent the transmission of HIV						
1	2	3	4	5	6	7
Supervision Area	**Sample Size n**	**Correct Responses c**	**Mini Proportion p=c/n**	**Population Size N**	**Weight Wt=N/∑N**	**Weight*Mini Percent Wt*p**
A	19	11	0.579	13,877	0.318	0.184
B	19	12	0.632	14,567	0.333	0.210
C	19	8	0.421	8,576	0.196	0.083
D	19	4	0.211	2,000	0.046	0.010
E	19	7	0.368	4,687	0.107	0.040
Totals	95	42	0.442	43,707		0.526

Notice that Columns 1–4 are exactly the same in this table as in the previous one. Column 5 has the same ∑N (43,707), but the individual SA populations are very different. Columns 6 and 7 are also different because they use the numbers from Column 5.

While our unweighted coverage estimate is still the same—44.2%, the weighted average here is very different—52.6%. Let us examine why. Notice that the two SAs with the highest number of correct responses—A and B (and, thus the highest mini-proportions) are also those with the largest populations. The SA with the lowest number of correct responses—SA D—has a very small population relative to them. Thus, SA D's lower mini-proportion is given less weight, and those with larger populations are given more weight. This acts to adjust upward the overall program coverage estimate.

We end this section—without any analysis—by merely demonstrating the formulas necessary to calculate 95% confidence intervals for these two weighted percentages. You can use the approaches provided here to carry out your own calculations of weighted percentages and attendant confidence intervals.

Let us examine the first example. We reprint the table for calculating the weighted percentage from above:

Percentage of women age 15–49 who know at least two ways to prevent the transmission of HIV						
1	2	3	4	5	6	7
Supervision Area	Sample Size n	Correct Responses c	Mini Proportion p=c/n	Population Size N	Weight Wt=N/∑N	Weight*Mini Percent Wt*p
A	19	11	0.579	10,718	0.245	0.142
B	19	12	0.632	6,379	0.146	0.092
C	19	8	0.421	9,379	0.215	0.090
D	19	4	0.211	9,731	0.223	0.047
E	19	7	0.368	7,500	0.172	0.063
Totals	95	42	0.442	43,707		0.435

8	9	10	11	12	13	14
Supervision Area	Mini Percent p=c/n	q=1-p	p*q	Population Size N	Weight Wt=N/∑N	Wt^2*((p*q)/n)
A	0.579	0.421	0.244	10,718	0.245	0.00077
B	0.632	0.368	0.233	6,379	0.146	0.00026
C	0.421	0.579	0.244	9,379	0.215	0.00059
D	0.211	0.789	0.166	9,731	0.223	0.00043
E	0.368	0.632	0.233	7,500	0.172	0.00036
Totals				43,707		0.00242

- Column 8 is merely a listing of the SAs.
- Column 9 is identical to column 4 above and is represented symbolically by “p.”
- Column 10 is needed to calculate what is known as the **“standard error”** (used to calculate the confidence interval) and is merely 1-p and we will represent it as “q”. In other words, for SA A, q=1-.579=.421.
- Column 11 multiplies these two numbers—p and q together (again, necessary to calculate the standard error). So for SA A, p*q is .579*.421=.244.
- Column 12 is the same as Column 5 in the previous table and is the population of each SA and the total program area population.
- Column 13 is the same as Column 6 in the previous table and is the weight of each SA.
- Finally, Column 14 calculates a “weighted standard error” by first squaring the weights from Column 13 (represented in the formula as $Wt^{2)}$ and then multiplying them by the product of p*q divided by n (taken from Column 2 of the previous table—which is 19 in each case). Again, the individual rows of this column is of no interest to us but only the total (sum) of each row. This sum is the standard error of the percentage and is plugged into the following formula to calculate the 95% confidence interval for the weighted proportion using this formula:

95% Confidence Interval for the Weighted Mean=$1.96 \pm \sqrt{(\sum(Wt^2 * ((p*q)/n))}$

Or, in this example: $.435\text{-}1.96*\sqrt{(0.00242)} = .338$ or 33.8%

$.434+1.96*\sqrt{(0.00242)} = .531$ or 53.1%

Thus, we are 95% confident that the true proportion of women age 15–49 in our program area who know two ways to prevent the spread of HIV lies between 33.8% and 53.1%

We conclude this section providing the same calculations for the second example:

Percentage of women age15–49 who know at least two ways to prevent the transmission of HIV						
1	**2**	**3**	**4**	**5**	**6**	**7**
Supervision Area	**Sample Size n**	**Correct Responses c**	**Mini-Proportion p=c/n**	**Population Size N**	**Weight Wt=N/∑N**	**Weight*Mini Percent Wt*p**
A	19	11	0.579	13,877	0.318	0.184
B	19	12	0.632	14,567	0.333	0.210
C	19	8	0.421	8,576	0.196	0.083
D	19	4	0.211	2,000	0.046	0.010
E	19	7	0.368	4,687	0.107	0.040
Totals	95	42	0.442	43,707		0.526

8	**9**	**10**	**11**	**12**	**13**	**14**
Supervision Area	**Mini Percent p=c/n**	**q=1-p**	**p*q**	**Population Size N**	**Weight Wt=N/∑N**	**Wt2*((p*q)/n)**
A	0.579	0.421	0.244	13,877	0.318	0.0013
B	0.632	0.368	0.233	14,567	0.333	0.0014
C	0.421	0.579	0.244	8,576	0.196	0.0005
D	0.211	0.789	0.166	2,000	0.046	0.0000
E	0.368	0.632	0.233	4,687	0.107	0.0001
Totals				43,707		0.0033

95% Confidence Interval for the Weighted Mean=$1.96 \pm \sqrt{(\sum(Wt^2 * ((p*q)/n)))}$

Or, in this example: $.526\text{-}1.96*\sqrt{(0.0033)} = .413$ or 41.3%

$.526+1.96*\sqrt{(0.0033)} = .639$ or 63.9%

Thus, we are 95% confident that the true proportion of women age 15–49 in our program area who know two ways to prevent the spread of HIV lies between 41.3% and 63.9%

PARALLEL SAMPLING FOR LQAS

Throughout all of our discussions of LQAS to this point, we have used a single indicator to illustrate how the approach works. We have done this to keep things simple. However, as we know, programs never concern a single issue and rarely have one group of interest in terms of health outcomes sought. Here we have used the example of women age 15–49 and the focus has been on their knowledge of ways to prevent the spread of HIV.

In most community-based health programs, we would probably be interested in other things like improving the health and nutritional status of children and practices of caregivers that are likely to affect those outcomes. Further, in order to assess changes in practice, we are likely to need to focus on subgroups of these such as children. For example, if we are concerned with infant and child feeding practices in a population we would want to assess children less than 6 months of age who should currently be breastfed or children 6–23 months who should be breastfed while other foods are being given. In general, we want to ask questions of caregivers who are currently practicing (or should be practicing) certain behaviors as opposed to asking them to recall what they did in the past when their child was younger. So, using the example of infant and child feeding, we want to ask mothers of children who are currently under 6 months of age about their breastfeeding practices rather than ask mothers whose children are older than 6 months of age (if we want to get at the issue of exclusive breastfeeding in children under 6 months of age).

As a result, when gathering information from respondents about their practices, we may randomly select a respondent who is not in a position to give us all the information we seek. We have selected a single interview site using the methods described previously but may find that the final respondent we select cannot respond to all the questions necessary to calculate the indicators of interest to our project. In such a case we need to find an approach to sampling at the interview site that allows us to find appropriate respondents for each indicator. Such an approach is referred to as parallel sampling, and we described the procedures for this in detail in this section. Before doing so, however, let us see an example of how this might work in practice.

Table M3.15 provides an example of the different response groups—referred to here as sample groups because we need to be sure that our sample includes one response for each group for the purposes of calculating the indicator given.

Table M3.15: Response groups

Sample groups	Example of indicators of interest in this group
Children 0–23 months	1. Percentage of mothers with children 0-23 months who received at least two Tetanus Toxoid doses before the birth of the youngest child
	2. Percentage of mothers of children age 0-23 months who live in a household with soap at the place for hand washing
Children 12–23 months	3. Percentage of children aged 12-23 months who received measles vaccine according to the vaccination card or mother's recall by the time of the survey
Children 0–5 months	4. Percentage of children 0-5 months who were exclusively given breast milk the day prior to the interview
Children 6–23 months	5. Percentage of children age 6-23 months fed according to a minimum of appropriate feeding practices
Children 0–23 months who had diarrhea in the past two weeks	6. Percentage of children age 0-23 months with diarrhea in the last 2 weeks who received oral rehydration solution (ORS) and/or recommended home fluids

Let us examine this table, which is not meant to be exhaustive but is representative of the challenge of having indicators for a program that require different respondent or sample groups. Suppose that the program seeking information about these indicators is focusing on increasing immunization coverage and on improving the nutritional status of children via improved feeding practices and better prevention and treatment of diarrhea.

Let's suppose through our sampling approach that we select a woman who has a seven-month-old child. Suppose further that the child has not had diarrhea in the past two weeks. In such a case, we could ask the mother questions related to indicators 1, 2, and 5 but we could not ask her questions related to the others. Although the child may have received some immunizations, for example, in most places measles immunizations are not given until after the ninth month. Also, if we want to promote exclusive breastfeeding in children under six months of age, we cannot ask about this because the respondent's child is over six months of age. Since the child did not have diarrhea in the past two weeks, we cannot ask about practices related to that event.

If we ended our work in this community after interviewing just this one respondent, we would not, over the course of interviews in the entire SA, have 19 respondents for each indicator and could neither assess whether the SA is reaching targets vis-à-vis those indicators nor could we produce a precise estimate of program wide-coverage because our total sample size for some variables could be far below 95 (making for very imprecise estimates).

Therefore, we need to have a sampling approach that enables us to find enough respondents—19 in each SA—for each indicator. To do this we use parallel sampling. Here are the steps to be followed.

1. Determine the number of sample groups.

To do this you must examine the denominators of all the indicators of interest and for which you will be collecting information in the survey. Organize indicators into groups of like denominators. Table M3.15 provides a simple example of what this might look like.

2. Develop questionnaires for each sample group.

In example in Table M3.15, you would need five different questionnaires. (A separate questionnaire is needed for each sample group.) For each sample group, create a questionnaire that has questions relevant to those indicators only.

In our example, you could start out by developing a questionnaire for mothers of children 0–23 months that has questions that can be asked of all mothers. In Table M3.15, this would mean including questions on hygiene and mother's immunization

Next, you would develop the questionnaire for 12–23 months and then continue with the other sample groups. Of course there are other items (age of mother, number of children, etc.) that you would want to collect from everyone, so make sure all of them are on each questionnaire. Make sure that each questionnaire has room to record other information, such as the interviewer and respondent identification information and consent information. Interviewers must fill out one set of questionnaires that includes all the sample groups for each interview site that is part of the survey. It is important for supervisors to organize the questionnaires so that the interview team that works in each interview has one questionnaire from each sample group to fill out.

3. Select interview sites.

We have already described this process in detail above and will not return to it here.

4. Select the first household.

Again, we have already discussed options for doing this and will not repeat them here.

Once you have identified the first household, we recommend the following approach—continuing with the same example we started with based on the indicators in Table M3.15).

5. Administer the questionnaires.

At the first household, ask if a child 0–23 months lives there and if the mother also lives there. If no child of this age range lives in the house move to the house with the nearest door to the door of the first house and again ask if a child 0–23 months lives there and if the mother also lives there. At the first household with a child in the appropriate age range, ask all the questions on the questionnaire for mothers of children 0–23 months. Then ask questions on the questionnaire pertaining to one of the other sample groups based on the age of that child. For example, if the first child is 13 months old, then the interviewer should fill out the questionnaire for mothers of children 0–23 months. Next, fill out the questionnaire for children 12–23 months. After that, ask if the child had diarrhea in the last two weeks. If the child had diarrhea, then fill out the appropriate illness questionnaire.

After finishing all relevant questionnaires for the first household, identify the questionnaires that still need to be completed (other sample groups). Proceed to the second household following the protocol for selecting second and subsequent households (see Module 2 on cluster sampling for options for selecting subsequent households) and ask if there is a child 0–23 months. If so, identify the age of the child and determine whether that child meets the criteria for any of the questionnaires that remain to be completed (either age-specific or because they had diarrhea in the past two weeks). Use the questionnaires that match the age or illness profile of the child, if any.

After this, continue on to the next household, using agreed upon protocols, if there are any questionnaires that still need to be completed. Continue this process until all questionnaires have been completed. This means that a child in each sample age group and with each illness has been identified and each questionnaire is completed.

If the interview site is very small and the interviewer has exhausted all the households in it, then move to the nearest community (not yet selected) and consider it an "extension" of the first interview site. Find the nearest household to the original interview site, and continue the interviews.

In the case of two appropriate children being present at the initial household, select one child selected randomly and interview the mother about that child using all the appropriate questionnaires. Afterward, the interviewer can interview the same mother about the second child using questionnaires for sample groups that were *not applied* to the first child. For example, if two children 0–23 months (one is 23 months and the other is four months) live in the first household, the interviewer randomly selects one child. If this child is the four-month-old, then the interviewer fills out questionnaires for 0–23 months and 0–5 months.

The interviewer then asks if this child was ill in the last two weeks. If the child was not ill, then the interviewer proceeds to ask about the 23-month-old child. The interviewer applies questionnaires for 6–23 and 12–23 months for the second child. If this child was ill with diarrhea in the last two weeks, then the interviewer fills out this questionnaire also.

Note that the application in one household of these questionnaires, *each one of which collects different information,* does not introduce bias into the sample. This is because the household was selected randomly as required by LQAS and because a mother was only asked once about questions for each indicator, even though some of these questions were asked about a child with one age and other questions were asked about a child of a different age. This would not be the case if two questionnaires about the same indicators (e.g., control of diarrhea) were applied in the same household for children of different ages. It is important to remember that two children from the same sample group cannot be interviewed in the same household. In the previous example, if both children had diarrhea during the previous 2 weeks, the diarrhea treatment questionnaire is applied only to one child.

In following this protocol, it is important to remember that it is necessary to obtain one answer to each question in each sample group questionnaire for each interview site. You then repeat this for each interview site in the supervision area. Because you sample 19 for each indicator, there will be a total of 19 answers to each question in each sample group questionnaire in each supervision area. If you have five supervision areas, this will result in 95 answers to each question in each sample group questionnaire for the entire project area.

SAMPLE SIZE ISSUES—IDEAL AND REAL

Up to this point we have used a consistent example of sample sizes of 19 for each SA in a case in which there are 5 SAs. Program wide this yields a total sample size of 95, which has the helpful feature of providing point estimates of coverage whose 95% confidence intervals are always approximately ±10 percentage points. However, the real world will probably not always be this simple, and sample sizes may have to be adjusted for at least two distinct reasons:

1. For whatever reason, even though you planned for 19 responses for each indicator per SA, you ended up with fewer than that. This leads to the question of what to do to assess whether an SA has reached a target given that you can no longer consult the LQAS table for a sample size of 19 that we have been using for a given indicator. While missing one or two responses in one or two SAs will not have much impact on the precision of your overall program coverage estimate (when you combine the results from all SAs), it will affect your ability to assess the results for a given SA vis-à-vis a target or the average coverage for the program area. We will examine ways to deal with this issue.

2. Another problem concerns the actual number of SAs you will have in a given program area. As noted at the outset, the choice of how many SAs to have should be based on real world needs, such the actual number of supervisors or program teams responsible for unique program zones. You should probably avoid simply creating SAs in order to end up with five (which has the benefits we discussed), but the question then becomes, what happens to overall program estimates if you have only three or four SAs? Keep in mind that having fewer than five SAs does not affect your ability to assess the individual SAs you have in relation to targets you set. However, having total sample sizes of 76 (for four SAs) or even 57 (for three SAs) will affect the precision of estimates for overall program coverage. We will examine options for dealing with these issues.

Let us examine the first challenge. It is bound to happen that in a given SA for a given indicator—due to problems in the field, lost data, or other errors—you will end up with fewer than 19 respondents for a given indicator. What can you do? The simple answer is to say that you can use an LQAS table with decision rules other than 19 (less than 19, in this case). If you have only 17 responses for a given indicator, for example, you can consult a table for decision rules for different coverage levels for this sample size. Obviously, with samples less than 19, there is some loss of precision—a greater likelihood of misclassification in one way or another. To see how this might work, we provide below an LQAS table that includes sample sizes below 19—from sample sizes 12–18. Let us examine this table in detail. You can, by now, understand how to find a decision rule for a given coverage level (percentage) for sample sizes other than 19. We can therefore proceed to a discussion of the loss of precision—the greater probability of misclassification—that occurs with smaller sample sizes. Examine the table first.

Table M3.16: LQAS table: Decision rules for sample sizes of 12–19 and coverage targets/average of 10–95%

Sample size*	Average coverage (baselines)/annual coverage target (monitoring and evaluation)																	
	10%	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**	**45%**	**50%**	**55%**	**60%**	**65%**	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**
12	N/A	N/A	1	1	2	2	3	4	5	5	6	7	7	8	8	9	10	11
13	N/A	N/A	1	1	2	3	3	4	5	6	6	7	8	8	9	10	11	11
14	N/A	N/A	1	1	2	3	4	4	5	6	7	8	8	9	10	11	11	12
15	N/A	N/A	1	2	2	3	4	5	6	6	7	8	9	10	10	11	12	13
16	N/A	N/A	1	2	2	3	4	5	6	7	8	9	9	10	11	12	13	14
17	N/A	N/A	1	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
18	N/A	N/A	1	2	2	3	5	6	7	8	9	10	11	11	12	13	14	16
19	N/A	N/A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

N/A: *not applicable*, meaning LQAS cannot be used in this assessment because the coverage is either too low or too high to assess an SA. This table assumes that the beta errors are the probability of misclassifying an SA that falls 30 percentage points below the target.

[light grey box]: shaded cells indicate where *alpha* or *beta* errors are ≥ 10%.

[dark grey box]: shaded cells indicate where *alpha* or *beta* errors are > 15%.

To fully exploit this table we have to formally introduce two concepts that we have discussed in other terms (and briefly in footnotes above)—the concept of alpha and beta errors. Let us use the sample size of 19 and coverage of 80% (decision rule of 13) to define these. As we have said before, if the true coverage percent for a given indicator is 80% in a population then with a sample size of 19, we are likely to find 13 respondents in our sample who respond correctly. In fact, 93% of the time we will find 13 or more and correctly classify the SA. The alpha error—which is merely the inverse of 93% or 7% in this case—is the probability of misclassifying an 80% percent population as not being 80% because we found fewer than 13 who answered correctly in the sample. The advantage of sample size 19 is that for all average coverage levels the alpha error is always below 10%. That is why in the above table none of the rows for sample size 19 is highlighted

Beta error is a bit different. LQAS is set up to correctly identify those who ARE reaching a coverage level more than 90% of the time and is also set up to identify at this same probability those SAs that fall *30 percentage points* below the coverage level. Thus, in the case of 80% coverage, the table is set up so that if the true coverage in the population is 50% or below we will correctly classify it as not having reached the target 92% of the time. In other words, we will find fewer than 13 who respond correctly 92% of the time. The 8% of the time when we find 13 or more when the true coverage is 50% is known as the beta error. Again, the advantage of 19 is that it keeps the beta error (defined in this way) below 10% for all coverage levels.

Now we can understand what the other rows on this table mean when they say that alpha or beta errors for a given sample size and coverage level are >10% or >15%. In other words, for smaller sample sizes, the probability of misclassification in either direction increases. This does not make LQAS less useful—especially for those sample sizes and coverage levels in the above table that are lightly shaded, indicating probabilities of misclassification as we have defined them of between 10 and 15 percent.

We therefore strongly recommend that you use this table, recognizing that for samples of less than 19, the probabilities of misclassification increase. You should note this and be cognizant of it but use the table to assess individual SAs. For specific details on the exact levels of alpha and beta error for each coverage level, please refer to Annex 4: Alpha and Beta Errors (n=19). The table in this annex provides exact alpha and beta errors for coverage levels 10-95% (and attendant decision rules) for LQA sample size 19 and provides a link to a document with the same information for samples size 10 to 30.

We now turn to the final issue raised in point two above—what to do when you have fewer than five SAs. The main issue we deal with in these cases is the loss of precision of overall program estimates because our total combined sample sizes are smaller. Loss of precision, quite simply, means that for a given point estimate (weighted percentage, for example), our confidence interval will be wider. The general recommendation we make is this: If you have three or four SAs, increase your sample size in each SA so that your total sample combining all SAs is 95 or more.

The following five example tables with eight, six, five, four, and three SAs illustrate this, using procedures we developed previously. Since confidence intervals are widest when coverage is

50%, we have used this as our example. Examine these three tables in terms of what they tell us about the precision of estimates. Notice that the confidence intervals increase as the number of SAs decreases. It will be up to your program to determine at what level the loss of precision becomes too great to bear.

The main point is to be very practical when defining the number of SAs. We reiterate the belief that each one should conform to a real management unit. This is important for the management reasons we have discussed throughout this section.

Table M3.17: Example 1: Eight SAs

SA	n	Corrects	mini %	N	wt	wt*(mini%)
1	19	7	0.368	10,718	0.16	0.06
2	19	14	0.737	6,379	0.09	0.07
3	19	5	0.263	9,379	0.14	0.04
4	19	16	0.842	9,731	0.14	0.12
5	19	6	0.316	7,500	0.11	0.04
6	19	9	0.474	8,000	0.12	0.06
7	19	10	0.526	7,500	0.11	0.06
8	19	10	0.526	8,000	0.12	0.06
	152	77		67,207	**Weighted Coverage**	**0.500**
Confidence Interval = Weighted Coverage Plus or Minus 0.073 or 0.427-0.573						

Table M3.18: Example 2: Six SAs

Cohort	n	Corrects	mini %	N	wt	wt*(mini%)
1	19	7	0.368	10,718	0.21	0.08
2	19	14	0.737	6,379	0.12	0.09
3	19	5	0.263	9,379	0.18	0.05
4	19	16	0.842	9,731	0.19	0.16
5	19	6	0.316	7,500	0.15	0.05
6	19	10	0.526	8,000	0.15	0.08
	114	58		51,707	**Weighted Coverage**	**0.501**
Confidence Interval = Weighted Coverage Plus or Minus 0.082 or 0.418-0.582						**0.082**

Table M3.19: Example 3: Five SAs

Cohort	n	Corrects	mini %	N	wt	wt*(mini%)
1	19	7	0.368	10,718	0.25	0.09
2	19	14	0.737	6,379	0.15	0.11
3	19	5	0.263	9,379	0.21	0.06
4	19	16	0.842	9,731	0.22	0.19
5	19	6	0.316	7,500	0.17	0.05
	95	48		43,707	**Weighted Coverage**	**0.50**
Confidence Interval = Weighted Coverage Plus or Minus 0.090 or 0.410-0.590						**0.090**

Table M3.20: Example 4: Four SAs

Cohort	n	Corrects	mini %	N	wt	wt*(mini%)
1	19	7	0.368	10,718	0.30	0.11
2	19	14	0.737	6,379	0.18	0.13
3	19	6	0.316	9,379	0.26	0.08
4	19	13	0.684	9,731	0.27	0.18
	76	40		36,207	**Weighted Coverage**	**0.50**
Confidence Interval = Weighted Coverage Plus or Minus 0.107 or 0.393-0.607						**0.107**

Table M3.21: Example 5: Three SAs

Cohort	n	Corrects	mini %	N	wt	wt*(mini%)
1	19	9	0.474	10,718	0.40	0.19
2	19	14	0.737	6,379	0.24	0.18
3	19	7	0.368	9,379	0.35	0.13
	57	30		26,476	**Weighted Coverage**	**0.50**
Confidence Interval = Weighted Coverage Plus or Minus 0.128 or 0.372-0.628						**0.128**

What then might we do in order to increase our sample sizes in the case in which we have four or even three supervision areas? One obvious option is to increase the sample size in each SA so that the total sample might approach a level for which the precision of a 95% confidence interval is 10 percentage points (a sample size of around 90 or so). For a program with four SAs, this implies taking samples in the range of 23–25 per SA. For a program with three SAs, this implies a sample of 30–32 in each SA. Obviously increasing sample sizes also increases the cost of data collection.

At the same time, these increases do not greatly decrease the alpha and beta errors that would enable us to more accurately classify individual SAs. Basically, increasing sample sizes in this way merely increases the precision of overall program coverage estimates. We say merely not because this is not valuable per se—it may be valuable to your program—but because increasing sample sizes in each SA does not have other advantages.

However, if you are increasing the sample size you cannot use the decision rule table for sample size 19 we have been using. Just as there are decision rule tables for sample sizes fewer than 19, such tables exist for sample sizes greater than 19. Table M3.22 includes sample sizes of coverage levels from 10–95% and sample sizes from 12–30 below.

The decision about which sample size to use is yours to make on the basis of such considerations as cost, time, and need for precision.

Table M3.22: LQAS Table: Decision rules for sample sizes of 12–30 and coverage targets/average of 10–95%

Sample size*	Average coverage (baselines) / annual coverage target (monitoring and evaluation)																	
	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
12	N/A	N/A	1	1	2	2	3	4	5	5	6	7	7	8	8	9	10	11
13	N/A	N/A	1	1	2	3	3	4	5	6	6	7	8	8	9	10	11	11
14	N/A	N/A	1	1	2	3	4	4	5	6	7	8	8	9	10	11	11	12
15	N/A	N/A	1	2	2	3	4	5	6	6	7	8	9	10	10	11	12	13
16	N/A	N/A	1	2	2	3	4	5	6	7	8	9	9	10	11	12	13	14
17	N/A	N/A	1	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
18	N/A	N/A	1	2	2	3	5	6	7	8	9	10	11	11	12	13	14	16
19	N/A	N/A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
20	N/A	N/A	1	2	3	4	5	6	7	8	9	11	12	13	14	15	16	17
21	N/A	N/A	1	2	3	4	5	6	8	9	10	11	12	13	14	16	17	18
22	N/A	N/A	1	2	3	4	5	7	8	9	10	12	13	14	15	16	18	19
23	N/A	N/A	1	2	3	4	6	7	8	10	11	12	13	14	16	17	18	20
24	N/A	N/A	1	2	3	4	6	7	9	10	11	13	14	15	16	18	19	21
25	N/A	1	2	2	4	5	6	8	9	10	12	13	14	16	17	18	20	21
26	N/A	1	2	3	4	5	6	8	9	11	12	14	15	16	18	19	21	22
27	N/A	1	2	3	4	5	7	8	10	11	13	14	15	17	18	20	21	23
28	N/A	1	2	3	4	5	7	8	10	12	13	15	16	18	19	21	22	24
29	N/A	1	2	3	4	5	7	9	10	12	13	15	17	18	20	21	23	25
30	N/A	1	2	3	4	5	7	9	11	12	14	16	17	19	20	22	24	26

N/A: not applicable, meaning LQAS cannot be used in this assessment because the coverage is either too low or too high to assess an SA. This table assumes the lower threshold is 30 percentage points below the upper threshold.

: shaded cells indicate where *alpha* or *beta* errors are ≥ 10%.

: shaded cells indicate where *alpha* or *beta* errors are > 15%.

CONCLUSION

While no single document can address all possible issues for all possibly types of population-based health survey, we have presented in simple language all the essential steps to the design and implementation of rapid surveys. The appendix provides a set of key resources with details on specific points which may have been treated only briefly in this handbook.

We hope this handbook will provide the readers with all the essential elements and links to appropriate resources to custom design their own surveys, adapted to their local information and management needs.

Many countries have seen local survey research groups develop the range of skills required. Even though some NGOs develop internal capacity for the conduct of surveys, many more will probably find it worthwhile to invest in outsourcing to these local survey groups. Even in this case, this handbook will provide a useful tool for supervision and quality assurance.

We have sought to present honestly what is reasonably achievable at project or district levels, with an honest treatment of constraints and limitations. In doing so we have detailed two main sampling approaches, which have been and are being used by many projects to analyze a (baseline) situation, define priorities, monitor progress, and evaluate the achievement of objectives. While questions will always remain about what more could be done—such as demonstrating impact through more advanced designs—the reality is that too many interventions are carried out without even the basic information provided by these rapid surveys. Improvements on the proposed approaches are always desirable, and we have discussed how certain options for sampling have been put forth in the literature. The real test—and this must be clear—for improving on rapid survey modalities, is not whether they can be brought closer to the precision and confidence of large surveys, but how any marginal improvement in quality and precision will translate in terms of transaction cost and value to local managers.

These small scale surveys do not pretend to replace large national surveys such as the DHS or MICS, which are considered gold standards for population health information at national and regional levels. However, because they require considerable investment of time, effort and budget, national surveys are usually only conducted every five years and do not provide project or district level information. Small scale rapid surveys fill in the information gap for local managers, who have to make time-bound decisions. Opposing or contrasting the two types of surveys emanates from a misguided thought process. Both are important and have their role.

The take home message to the reader is:

- Rapid health surveys have an essential role to play in guiding strategies and programs at the local level.
- They are feasible, (relatively) quick and inexpensive, reliable and valid within the parameters of their appropriate use.
- Similar to any other survey designs or data collection methods, they require a great level of attention at every step of design and implementation to maintain their validity.
- Before engaging in expensive (and largely futile) efforts to (retrospectively) try to show impact, district health managers and project managers should be empowered to establish

the trail of evidence which rapid surveys can provide, from the onset of new efforts and on a regular basis.

We hope this manual will encourage this process.

ANNEX 1: KEY RESOURCES

This annex provides references for professionals who want more details about specific topics. This is not an exhaustive list of all references. However, it contains references that we found useful for the preparation of this handbook and that we feel would be useful to professionals who have more advanced questions about surveys.

The following are websites that we recommend if you need more detailed information. See specifics of each in the table below.

1. The Maternal and Child Health Integrated Program (MCHIP) Non-Governmental Organization (NGO) strengthening site: www.MCHIPNGO.net or www.childsurvival.com
2. The CORE Group: www.coregroup.org
3. The Demographic Health Survey (DHS) site: www.measuredhs.com
4. MEASURE Evaluation: www.cpc.unc.edu/measure
5. The Research Methods Knowledge Base: http://www.socialresearchmethods.net/kb
6. The Flexible Fund: www.flexfund.org
7. Food and Nutrition Technical Assistance (FANTA 2), www.fanta-2.org
8. Epi Info: http://www.cdc.gov/epiinfo/
9. UCLA Department of Epidemiology, Rapid Survey Course: http://www.ph.ucla.edu/epi/rapidsurveys/RScourse/RSrapidsurveys.html
10. Malaria Indicator Survey: Basic Documentation For Survey Design And Implementation: http://www.searo.who.int/EN/Section10/Section21/Section1365_11100.htm

In these tables we mostly provide information on documents and resources available at the above-referenced websites. In several cases we provide links to specific documents while noting that each of the websites undergoes frequent modifications and updates and, therefore, we cannot ensure that any of the links listed here will remain in place. They were all active as of December 2009.

Useful documents	**Key points about documents**
MCHIP NGO Strengthening	
Knowledge Practice and Coverage Resources (KPC) Many resources available through this link: http://www.childsurvival.com/kpc2000/kpc2000_new_summary.cfm The "Field Guide" is available at: http://www.childsurvival.com/kpc2000/FieldGuide_Sept03.pdf The Methodology and Sampling Issues for KPC Guide is available at: http://www.childsurvival.com/kpc2000/method.pdf	• Rapid small population survey developed for Child Survival and Health Grants Program (CSHGP) that is regularly updated to be consistent with state-of-the-art indicators and other important evaluation sources (i.e., DHS, MICS, Malaria MERG, and USAID) • Modules (questionnaires, tabulation plans and indicator definitions) for technical areas: Water and Sanitation; Breastfeeding and Infant and Young Child Feeding; Immunization; Anthropometrics; Control of Diarrhea; Acute Respiratory Infections; Malaria; Maternal and Newborn Care; Child Spacing and HIV/STIs. • Methodology and Sampling Guide • Field Guide • Rapid CATCH module: Questionnaire and Tabulation plan for a reduced set of standard indicators from the technical areas. • Implemented at sub-national or district level in rural and urban areas of Africa, Asia/Near East, Latin America, and Europe and Eurasia
Monitoring and Evaluation Technical Reference Material (M&E TRM) http://207.226.255.123/working_groups/ME_TRMs_2006.pdf	This document was prepared for CSHGP. It contains: • Basic M&E concepts • Process for developing an M&E plan that is linked to program design • Conceptual frameworks • Data sources and collection methodologies

Useful documents	Key points about documents
CORE Group	
KPC Trainer of Survey Trainers (TOAST)[1] http://207.226.255.123/working_groups/kpc_training/welcome.html	• Curriculum to teach how to implement a KPC survey. • Contains three modules: (1) Training the Core Team; (2) Training Supervisors and Interviewers; and (3) Training the Post-Survey Analysis Team. • Topics include: Staffing patterns and roles; Developing a questionnaire; Sampling Terminology; 30x10-cluster sampling; LQAS sampling; Household selection; Logistics Planning; Data Analysis; and Report Writing.
Assessing Community Health Programs A Trainer's Guide Using LQAS for Baseline Surveys and Regular Monitoring Trainer's guide and participant handouts available (respectively) at: http://207.226.255.123/working_groups/lqas_train.html and http://207.226.255.123/working_groups/LQAS_Participant_Manual_L.pdf	• Contains basic step by step process for understanding LQAS concepts; Implementing surveys using LQAS; and Analyzing results using a decision table.
Lot Quality Assurance Sampling (LQAS) Protocol for Parallel Sampling http://207.226.255.123/working_groups/LQAS_Protocol_for_Parallel_Sampling.pdf	• Practical instructions for parallel sampling with LQAS. Developed from a technical advisory meeting on the subject.
LQAS Frequently asked Questions http://207.226.255.123/working_groups/LQAS_FAQ.pdf	• Tips for common concerns developed from a technical advisory meeting.

[1] Monitoring and Evaluation Working Group, CORE Group, *Knowledge, Practice, Coverage Survey Training Curriculum*, Washington D.C: December 2004.

Useful documents	Key points about documents
DHS	
Key Indicator Survey (KIS) Questionnaire and Guidance http://www.measuredhs.com/aboutsurveys/kis.cfm	• Provides monitoring and evaluation data for population and health activities in small areas—regions, districts, catchment areas—that may be targeted by an individual project, although they can be used in nationally representative surveys as well. • The KIS tool includes six questionnaires-a common Household Questionnaire and five individual questionnaires on Family Planning; Maternal Health; Child Health; HIV/AIDS; and Infectious Diseases. • Guidance manuals for KIS: Introduction; interviewer's manual; Sampling guidelines; Tabulation plan; Questionnaire user's guide; Manual on taking anthropometric measurements.
Demographic and Health Surveys (DHS) http://www.measuredhs.com/aboutsurveys/dhs/start.cfm	• Demographic and Health Surveys (DHS) are nationally representative household surveys that provide data for a wide range of monitoring and impact evaluation indicators in the areas of population, health, and nutrition. • Standard DHS Surveys have large sample sizes (usually between 5,000 and 30,000 households) and typically are conducted every 5 years, to allow comparisons over time. • Questionnaires, manuals, and guides are on the website. • Implemented in Africa, Asia/Near East, Latin America, and Europe and Eurasia
AIDS Indicator Survey (AIS) http://www.measuredhs.com/aboutsurveys/ais/start.cfm	• Provide countries with a standardized tool to obtain indicators for the effective monitoring of national HIV/AIDS programs. • Website contains methodology description, questionnaires and manuals.

Useful documents	Key points about documents
MEASURE Evaluation	
MEASURE Evaluation has developed a variety of monitoring and evaluation tools and guides. The following are important examples.	
Sample Vital Registration with Verbal Autopsy (SAVVY) http://www.cpc.unc.edu/measure/tools/monitoring-evaluation-systems/savvy	• SAVVY is a resource library of methods for strengthening vital events monitoring and measurement, including causes of death. • Manuals on the website include the following: data processing; budgets; interviewer, field officer, supervisor and supervisor coordinator manuals; Census household questionnaire; and Verbal Autopsy Certifier and Coding Manual.
Priorities for Local AIDS Control Efforts (PLACE) http://www.cpc.unc.edu/measure/tools/hiv-aids/place	• PLACE is a rapid assessment tool to monitor and improve AIDS prevention program coverage in areas where HIV transmission is most likely to occur. • Website contains the manual with step by step instructions for implementing PLACE; Instructions on identifying venues where people meet new partners; Interviewer guides; Questionnaires; Confidentiality agreement; Sample budget; and Report template.
MEASURE Evaluation. A Guide for Monitoring and Evaluating Child Health Programs[2] http://www.cpc.unc.edu/measure/publications/pdf/ms-05-15.pdf or http://207.226.255.123/working_groups/ms-05-15.pdf	Guide contains: • Program concepts of inputs, processes, outputs, and outcomes; • Value of a conceptual framework; • Indicator selection; • Data Sources; • Specific indicators for technical areas[3] of prevention of mother-to-child transmission of HIV (PMTCT); Newborn Health; Immunization; Integrated disease surveillance and response; Integrated management of childhood illness; Diarrhea, Acute Respiratory Infections; Fever; Growth Monitoring and Nutrition; and Mortality

[2] Anastacia J. Gage, Disha Ali, Chiho Suzuki; *A Guide for Monitoring and Evaluating Child Health Programs*; USAID, MEASURE Evaluation, World Health Organization, UNICEF, World Bank; September 2005.
[3] Because this document was published in 2005, some indicators may not be the most current versions.

Useful documents	Key points about documents
The Research Methods Knowledge Base[4]	
The Research Methods Knowledge Base is a comprehensive web-based textbook that addresses all of the topics in a typical introductory undergraduate or graduate course in social research methods. It covers the entire research process including: formulating research questions; sampling (probability and non-probability); measurement (surveys, scaling, qualitative, unobtrusive); research design (experimental and quasi-experimental); data analysis; and, writing the research paper. It also addresses the major theoretical and philosophical underpinnings of research including: the idea of validity in research; reliability of measures; and ethics.	
The Flexible Fund	
Program Design Monitoring and Evaluation (PDME) curriculum http://www.flexfund.org/resources/training/pdme.cfm	• The PDME course gives mid and senior level country managers the opportunity to acquire skills to develop project designs and monitoring and evaluation plans that are linked to these designs. • During this course, participants learn a 6-step process for developing a project design using a results framework and for developing a monitoring and evaluation (M&E) plan, both based on a situational analysis and an organized process for extracting and analyzing this information.

[4] Trochim, William M. *The Research Methods Knowledge Base, 2nd Edition*. Internet WWW page, at URL: <http://www.socialresearchmethods.net/kb/> (version current as of October 20, 2006).

Useful documents	Key points about documents
Food and Nutrition Technical Assistance (FANTA 2)	
LAYERS http://www.fantaproject.org/about/layers.shtml	• Layers is a software application that uses a hand-held computer to collect and analyze data based on principles of LQAS. Once the data is entered and analyzed, Layers generates automated reports. • Modules are available for Maternal Child Health and Nutrition; Food for Education; Food For Work programs; and commodities (from food aid to vaccines and drugs) • User manuals are available that can be loaded directly onto a handheld PDA. • Website contains a sample size calculator for LQAS using different sample sizes including SAs with less than 19 responses.
Alternative Sampling Designs for Emergency Settings: A Guide for Survey Planning, Data Collection and Analysis[5] http://www.fantaproject.org/downloads/pdfs/ASG_FINAL_Sept24.pdf	This is an extensive guide produced by the Fanta-2 Project at Academy for Educational Development (AED). Topics of interest include: • Sampling approaches • Comparing alternative cluster designs of 33x6, 67x3 and 30x30 • Using LQAS to measure acute malnutrition • Choosing the most appropriate sampling design • Questionnaire development • Data analysis and reporting
Methodology: Precision, time, and cost: a comparison of three sampling designs in an emergency setting[6] http://www.ete-online.com/content/5/1/6	• This article[7] describes a study done to compare precision, time and cost of 30x30 cluster surveys to two alternate sampling designs: 33x6-cluster and 67x3-cluster. • The article discusses advantages and disadvantages for different types of indicators and how this is influenced by different levels of intra-cluster correlation. LQAS is discussed in the context of 67x3 design

[5] Fanta-2 Project, Academy for Educational Development (AED), September 2009
[6] Megan Deitchler, Hedwig Deconinck, and Gilles Bergeron; published May 2 2008; Emergind Themes in Epidemiology 2008, 5:6 doi:10.1186/1742-7622-5-6 2008 Deitchler et al; licensee BioMed Cen tral Ltd.
[7] This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0). The article is available at http://www.ete-online.com/content/5/1/6.

Useful documents	Key points about documents
Epi Info	
Epi Info is a free software package developed by CDC for use by public health and medical professionals to develop questionnaires, customize data entry process, and enter and analyze data. • Modules include: data entry; analysis; analysis of nutritional data (Anthropometric data – nutchildren);	
UCLA Department of Epidemiology Rapid Survey Course[8]	
See especially: http://www.ph.ucla.edu/epi/rapidsurveys/RScourse/RSsoftware.html http://www.ph.ucla.edu/epi/rapidsurveys/RScourse/RSstmanual.html http://www.ph.ucla.edu/epi/rapidsurveys/RScourse/weights_clustersurveys.pdf	This site provides a full course in "rapid surveys. The site is maintained by Professor Ralph Frerichs to support rapid surveys. The site contains rapid survey software and publications on surveys and survey methodology. Of most use in relation to the current manual are the parts of the site that deal with available software with tutorials on how to use it for such things as the calculation of design effect and instructions on calculating cluster weights.

[8] UCLA Department of Epidemiology Rapid Survey Course created and managed by Ralph R. Frerichs

Useful documents	Key points about documents
Malaria Indicator Survey (MIS): Basic Documentation For Survey Design And Implementation	
Useful Documents	**Key points about documents**
This site provides basic documentation on how to conduct a malaria indicator survey. The methodology is taken from the DHS approach and presented in useful modules.	The MIS was developed by the Monitoring and Evaluation Working Group (MERG) of Roll Back Malaria. (www.rollbackmalaria.org). DHS has been a major contributor to its development. It is a stand-alone household survey to collect national and regional or provincial data from a representative sample of respondents. The survey is designed to help national malaria control programs and international health organizations with malaria programming. Key modules available at this time include (bolded indicates that they were referred to in the current manual: Overview of the MIS Documentation MIS Components– April 2005 Core Component 1: Household Questionnaire Core Component 2: Women's Questionnaire Core Component 3: Rationale Core Component 4: Interviewer's Manual **Core Component 5: Supervisor's Manual** Core Component 6: Guidelines for Interviewer Training **Core Component 7: Household Listing Manual** **Core Component 8: Sampling Guidelines** Core Component 9: Tabulations for Key Malaria Indicators Complementary Documents 1: A Field Guide to GPS Data Collection & GPS Cluster Position Form Complementary Documents 2: PDA Manual for Field Data Collection and Sampling **Complementary Documents 3: Calculating the Cost of the Malaria Indicator Survey**

The following resource is available for purchase:

Useful documents	**Key points about documents**
Survey Methodology: Wiley Series in Survey Methodology[9] Information about documents in this series (available for purchase) available at: http://www.wiley.com/WileyCDA/Section/id-300613.html	This is a comprehensive reference guide to survey methodology. Chapters are: • Introduction to survey methodology • Inference and error in surveys • Target populations, sampling frames and coverage error • Sample Design and Sampling Error • Methods of data collection • Non-response in sample surveys • Questions and answers in surveys • Evaluation survey questions • Survey interviewing • Post-collection processing of survey data • Principles and practices related to scientific integrity • FAQs about survey methodology

[9] Robert M. Groves, Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. *Survey Methodology*. Wiley Series in Survey Methodology. 2004 John Wiley and Sons, Inc.

Useful documents	Key points about documents
Qualitative Research for Improved Health Programs, a Guide to Manuals for Qualitative and Participatory Research on Child Health, Nutrition, and Reproductive Health[10] Available at: http://sara.aed.org/publications/cross_cutting/qualitative/qualitative.pdf	Guide to qualitative research manuals
Monitoring and Evaluation Manual by Medical Teams International This and other planning resources by Medical Teams International are available at: http://www.medicalteams.org/sf/learning_zone/learning_zone_cross_cutting.aspx	This manual is designed to help project managers, monitoring and evaluation coordinators, and project implementers gain familiarity with basic monitoring and evaluation concepts. The manual includes broad overview information, as well as details regarding how to make monitoring and evaluation plans, how to carry out specific monitoring and evaluation methods, such as surveys and qualitative research, and how to provide feedback of the results to communities. Considerable emphasis is given to preferred tools, surveys, and approaches used by Medical Teams International (MTI), most of which are standard for U.S.-based NGOs.

[10] Peter J. Winch, Jennifer A. Wagman, Rebecca A. Malouin,Garret L Mehl. Prepared by Department of International Health Johns Hopkins University, School of Hygiene and Public Health for Support for Analysis and Research in Africa (SARA) Project of Academy for Educational Development funded by U.S. Agency for International Development, Africa Bureau. *Qualitative Research for Improved Health Programs, a Guide to Manuals for Qualitative and Participatory Research on Child Health, Nutrition, and Reproductive Health*. January 2000.

ANNEX 2: CALCULATING THE COST OF A RAPID SURVEY

This annex is adapted from the KPC Training Guide. You can find the full reference for and a link to KPC training materials in Annex 1.

Rapid Survey Logistics and Management Planning Form

Choose survey dates. Mark on the form below the months in which other events might make it difficult to conduct the survey. Then choose the times that are least problematic.

Scheduling issues to consider	**Months of the Year (Western Calendar)**											
	J	**F**	**M**	**A**	**M**	**J**	**J**	**A**	**S**	**O**	**N**	**D**
Holidays (Tet, Christmas, Ramadan, etc.)												
Bad weather (monsoon, very hot weather)												
Times when respondents are less available (harvest/planting season, migration)												
Other project activities or scheduling conflicts for key persons (e.g., staff retreats)												
Disease prevalence patterns												
Food security and eating patterns (hungry season)												
Other issues												

Survey Dates: The survey will be conducted over approximately 28 days beginning on: __________________ and ending on _________________.

Enter your dates in calendar format and begin charting major activities.

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday

Mark dates for:
V = vehicles needed
CT = Core Team
S = Supervisors
I = Interviewers

Include dates for:
Field Testing
Training of Core Team, Supervisors, Interviewers and Post-Survey Team
Data Collection
Data Entry
Analysis Workshop

Personnel recruitment plan (drivers, data entry personnel and other support people will be included in other sections).			
Task:	**Survey Trainer Consultant**	**Supervisors**	**Interviewers**
How many?			
Who will recruit them? (Add to Action Plan)			
When will they be needed? (Calendar)			
From where will you recruit them? (Your PVO, partners, MOH, universities, or research institutions)			
Lodging arrangements			
Food arrangements			

Data Collection: If possible, arrange for the data collection phase to be conducted over 2 to 5 days. To achieve this, you need to identify enough people to form Survey Teams that can complete the 30-cluster survey within 2 to 5 days if you are using cluster sampling. In order to know how many person-days are needed to complete the data collection phase, determine the anticipated average length of the interview and the average walking distance between homes. Use this information to estimate the number of person-hours it will take to complete one cluster, and then extrapolate this information to determine how many person-hours it will take to complete the entire survey.

For example, if the survey takes 30 minutes to administer, and if walking time is about 10 minutes between households with children under 2 years of age in a given cluster, then one can expect an entire cluster of ten households to take ((30 minutes x 10 households) + (10 minutes x 9 households)) = 300 + 90 = 390 minutes / 60 = 6.5 person-hours. Therefore, with two people working the cluster, the cluster can be completed in 3.25 hours, or easily in one morning.

This is the usual pattern in areas where population dispersion is not very high, communities are within one hour or less travel time, and the interview time is close to 30 minutes: two Interviewers and one Supervisor can do one cluster in the morning and one cluster in the afternoon.

The number of teams needed depends in part on the <u>terrain</u> of the project area and the availability of transport during the survey. With a sample size of 300, if one team can complete 10 interviews in each of two clusters per day, then five teams can complete the survey in three days. If each team can complete only one cluster per day, then the (n=300) survey will take six days.

Transportation Plan				
Task:	**Vehicles**	**Drivers**	**Fuel and maintenance**	**Maps of survey sites**
What kinds/how many/ how much?				
When will you need them? (Calendar) (Include visits to villages to conduct mapping and brief community)				
Who will get these?				
From where will you obtain them? (PVO, partner, MOH, private hire)			Where will you buy fuel? Who will maintain vehicles?	
Other comments:				

Editing/Printing/Copying/Survey Forms and Other Materials

Estimation Matrix for Document Copying Needs			
Document	**A: Number of (1-sided) pages per document**	**B: Number of copies of each document needed**	**Total number of copies to be made (A x B)**
Surveys (for training and field use)			
Other training materials			
Field documents (maps or quick reference sheets)			
Tabulation tables			
Charts/graphs for presentations			
Rapid Survey Quality Checklists			
Rapid Survey report			
Other materials to disseminate results to the community and other stakeholders			
		Total of copies to be made:	

A 45-question *Rapid Survey Questionnaire* is about ten pages long. You need an additional 10% (in case one questionnaire in 10 needs to be recopied). In addition, there are five questionnaires per person for the rapid survey training (for the practice sessions). Assuming a sample size of 300 and that 20 people are trained, you need to make: (10 pages x 300 x 1.10) + (10 x 5 x 20) = 4,300 photocopy pages (2,150 sheets if copied double-sided).

For manual tabulation tables, you need about 100 pages for a full set. An additional set of tables should be kept as a record for future surveys. Some tables will also be needed in training activities (approximately 20%). Therefore: 100 x 2.2 = 220 copied sheets.

The *survey report* will probably be approximately 60 pages x 10 copies per feedback session x 2 feedback sessions + 10 extra photocopies of the report = (60 x 10 x 2) + (60 x 10) = 1,800 photocopied pages.

Total *photocopies* needed = 6,320. Typically, an office photocopy machine cannot usually handle this volume of photocopies and other arrangements may be necessary.

Lesson learned: Use a professional photocopy facility and contact and be prepared to use a backup facility in case of problems. Let the professional facility know your schedule and your photocopy needs in advance in case they need to purchase toner, etc.

Plan for Editing, Printing, and Copying Documents				
Task	**Who is responsible for this?***	**How many/how much?**	**When? (Calendar)**	**Where?**
Use computer software to edit the questionnaire file				
Locate good quality printer				
Identify a qualified person to do the editing				
Contact a professional photocopy facility **and** a back-up facility				Primary facility
				Back-up facility
Make copies of the questionnaire for the pre-test and training				
Make copies of the other materials for the pre-test and training				
Translate the questionnaire				
Make copies of the modified questionnaire shortly before the survey				
Other comments:				

* Add to Action Plan

Tabulation and Analysis of Data

Plan for Computerized Tabulation of the Survey				
Task	**Who is responsible for this?***	**How many/how much?**	**When? (Dates)**	**Where?**
Obtain computers, printers, paper, ink, and database software				
Identify trained computer personnel for data entry				
Identify personnel for supervising data entry, cleaning data and maintenance of computers				
Ensure alternative power sources if power is not reliable				
Other comments:				

* Add to Action Plan

Personnel Budget (Refer to the forms previously filled out for data)						
Type	**Number of persons**	**Number of days**	**Daily salary**	**Daily food/per diem cost**	**Daily lodging cost**	**Total cost***
Supervisors						
Interviewers						
Core Team members						
Rapid Survey Trainer Consultant						
Drivers						
Document preparation and copying support person(s)						
Data entry and computer maintenance personnel (if needed)						
Sub-total for personnel						

* Total Cost = (Number of Persons) x (Number of Days) x [(Daily Salary) + (Daily food/per diem) + (Daily Lodging)

Note: These forms are optimally incorporated within a spreadsheet like MS Excel.

Transportation Budget (Refer to the forms previously filled out for data)				
Item				
Type of vehicle	**Number of vehicles**	**Quantity**	**Cost per unit**	**Total cost**
		"Days"		
Diesel Fuel/ Petrol		Litres/gallons		
Maintenance (estimated labor and spare parts for vehicle maintenance and minor repairs needed to keep the survey on schedule)				
Subtotal for transportation				

Budget for Various Services and Equipment (Refer to the forms previously filled out for data)			
Service/Item	**Quantity**	**Cost per unit**	**Total Cost**
Equipment: Computers, printers, generators, software, etc.			
Translation			
Document copying			
Rental of office and meeting space			
Food costs during workshops			
Subtotal			

Summary of Estimated Budget	
Budget Category	**Subtotal**
Personnel	
Transportation	
Various Services and Equipment	
Estimated Grand Total:	

ANNEX 3: CALCULATING SAMPLES SIZES FOR VARIOUS SCENARIOS: FORMULAE AND EXAMPLES

This annex is adapted from the *Methodology and Sampling Issues for KPC Surveys.* You can find the full reference for and a link to this document in Annex 1: Key Resources.

In this annex we demonstrate sample size calculations for various survey objectives. We will use measles vaccination coverage in children 12-23 months of age as an example of an indicator we are monitoring:

- Example 1: Sample size determination for a baseline survey
- Example 2: Assessing achievement of objectives in the entire district
- Example 3: Analyses based on a final coverage estimate
- Example 4: Calculating a sample size for the comparison between two groups
- Example 5: Assessing the significance of an observed change

Example 1: Sample size determination for baseline survey

From the information available in other regions, we expect measles vaccination coverage to be around 40%, but we would like to assess the coverage in our region of intervention within10% of our estimate.

If we used simple random sampling to estimate this coverage, the appropriate formula for the sample size would be:

$$N = Z\alpha^2 \, pq / d^2 \qquad \textbf{(1)}^{(*)}$$

(*) Refer to appropriate reference for a discussion of the formula.

We can set the formula values as follows:

- $Z\alpha$=1.96 corresponding to a confidence level of 95%
- p = 0.4 (our expected coverage)
- q = 1-0.4 = 0.6
- d = accuracy desired = 10% = 0.10

We obtain:

$$N = (1.96)^2 \times 0.4 \times 0.6 / (0.10)^2 = 92$$

For reasons of economy, time and logistics, we have decided to use a 30-cluster sampling method to conduct our survey. As we have seen, this introduces a design effect (deff) in the precision of our estimate. A measles vaccination cluster survey in a neighboring region obtained a deff of 1.8, slightly lower than the value of 2.0 usually used to calculate cluster-survey sample size.
We can now correct the sample size needed in our cluster survey (N_c) to achieve the same level of precision of 10% by using the formula:

$$N_c = N * \text{deff} \qquad \textbf{(2)}$$

In this case,

$$N_c = 92 \times 1.8 = 166$$

This is, of course, the sample size that would be needed in the age group concerned by measles vaccination. As the KPC survey targets children 0 to 23 months of age and not only 12 to 23 months, we need to obtain a total sample size large enough to include 166 children in the 12 to 23 months of age sub-group. If we estimate (from available demographic data) that 45% of the sample of children 0 to 23 months of age will be in the target age range for this indicator of 12-23 months of age, different options are available to ensure this result:

1. We can increase the total sample size proportionally to our need for 166 children aged 12-23 months.

$$N_t = 166 \times (100/45) = 369 \qquad \textbf{(3)}$$

Where N_t is the total sample size.

In this case, we expect to have an appropriate sample size for our immunization coverage question.

2. As described in table 3 of section 3, we could also specifically oversample children aged 12-23 months so that 166 are sampled.
   This approach is quite simple and cost-effective.

3. If it appeared important to ask one series of question for children 0-11 months and another for 12-23 months, a parallel sampling strategy might be used. In this case, two different sample sizes should be calculated and we would use N_c =166 for the immunization coverage question.

4. If we could decrease the homogeneity within each cluster, either by stratifying by age group, or by improving the recruitment process (selection of each third or fifth household after the first one, initiating the randomization from different quadrants of the villages/clusters) we would decrease the deff of our survey. This approach is in fact feasible can only be established by experimentation in similar settings, about similar questions, and analysis with a computer software such as EPINFO.

Assuming we expect to have a lower deff, for example 1.2 instead of 1.8, we would then need a sample size of:

$N_c = 92 \times 1.2 = 111$ (see (2))

$N_t = (92 \times 1.2) \times (100/45) = 246$ (see (3))

Example 2: Assessing achievement of objectives in the entire district

We now have to determine the sample size for the final survey in order to assess whether a target measles vaccination coverage of 70% has been reached.

With the CSAMPLE program in EPI-INFO for the baseline survey (with 30 clusters where every 3rd household was selected), we found a deff of 1.5. We should assume the same deff for our final survey if we follow the same method.

If we used simple random sampling with

- $Z\alpha$=1.96 corresponding to a confidence level of 95%
- p = 0.7 (our target coverage)
- q = 1-0.7 = 0.3
- d = accuracy desired = 10% = 0.10

We would need:

$$N = (1.96)^2 \times 0.7 \times 0.3 / (0.10)^2 = 84$$

Using a 30 cluster sample, we obtain:

$$N_c = 84 \times 1.5 = 126$$

Using the same logic as for the baseline we would need a total sample Nt:

$$N_t = 126 \times (100/45) = 280$$

Or we could also simply over-sample in the 12-23 months old age group (see Example 1).

Example 3: Analyses based on a final coverage estimate

Let us now assume that 137 children aged 12-23 months were in our final survey. If 78 of them (or 56.9%) have been vaccinated against measles, we can use the CSAMPLE program in EPI-INFO, to obtain a 'correct' 95% confidence interval (as opposed to calculating an 'incorrect' confidence interval by ignoring the deff introduced by the cluster design). We obtain a deff of 1.06 and a 95% confidence interval of our estimate between 48.4% and 65.5%.

- **We conclude that we have failed to reach our target coverage of 70%.**

If we only 'recruited' 97 children aged 12-23 months in our final survey and 56 of them (or 57.7 %) have been vaccinated against measles. With a larger deff of 1.70, we would obtain a 95% confidence interval of our estimate between 44.7% and 70.7%.

- **We cannot conclude (statistically) that we have failed to reach our target coverage of 70%, at the 95% confidence level. But our best estimate is that we are 12% below our objective. Accepting a smaller confidence level (90% for example) we could, however, probably reject having reached our target, since its value is close to the margins of our 95% confidence interval. But it would have been more satisfactory to increase the sample size in the age-group and to decrease the deff, in order to be able to answer conclusively at the traditional 95% confidence level.**

With 59 children immunized out of 65 (90.8 %), and a deff of 1.38, we would obtain a 95% confidence interval of our estimate between 82.5% and 99.0%.

- **We conclude that we have reached our target coverage of 70% and are even above an 80% target, with a 95% confidence.**

Inversely, in spite of a high estimate on a larger sample (79.4% or 104 children out of 131), a large deff (e.g. 2.16) related to a high level of clustering of immunization, would yield a confidence interval between 69% and 89.6%, and would not allow us to conclude statistically that our estimate is statistically significantly superior to our preset objective. As a manager, we would report that our best estimate is that the region of intervention has reached 79% coverage and that there is no statistical evidence against the program having reached its objective.

Example 4: Calculating a sample size for the purpose of a comparison between two phases or two groups

Let us continue our example, using the 30-cluster method to conduct our surveys. Let us assume that a survey in a neighboring region using a 30-cluster design obtained a deff of 1.8.
One formula given for calculating the two sample sizes is given by:

$$N_1 = N_2 = \left[Z\alpha_{/2}\sqrt{[2pq]} + Z\beta\sqrt{[p_1q_1 + p_2q_2]}\right]^2/(p_1- p_2)^2 \qquad \textbf{(4)}$$

Where:

- N_1 = baseline sample size
- N_2 = final evaluation sample size
- $Z\alpha_{/2}$ is the Z value corresponding to the chosen level of risk α. ($Z\alpha_{/2}$ should be used in two-sided tests, and $Z\alpha$ should be used in one-sided tests.)
- $Z\beta$ is the Z value corresponding to the chosen level of risk β (it directly relates to the 'power' of the test as power = 1 - β); ($Z\beta$ = 1.28 for a power of .9)
- p_1 is the expected coverage at baseline
- $q_1 = 1 - p_1$
- p_2 is the expected final coverage
- $q_2 = 1 – p_2$
- $p = (N_1 \; p_1 + N_2 \, p_2) / (N_1 + N_2)$
- $q = 1 - p$

In fact more precise statistical software use a correction of formula (4), as follows:

$$N_1' = N_2' = N_1 \times \left[1 + \sqrt{\left(1+4(p_1- p_2)\right)}\right]^2 / 4(p_1- p_2)^2 \qquad \textbf{(5)}$$

If the expected coverage at baseline is 40%, and we want to be able to demonstrate an increase of 20 percentage-points in the final evaluation (meaning that we want to be able to demonstrate an increase from 40% to 60%), the sample size for each survey would be:

$$N_1 = N_2 = \{1.96\sqrt{[2(.5)(.5)]} + 1.28\sqrt{[(.4)(.6) +(.6)(.4)]}\}^2/(.4-.6)^2$$
$$N_1 = N_2 = 129$$

A simplified formula is available, and would yield a similar result:

$$N_1 = N_2 = [Z\alpha_{/2} + Z\beta]^2 \, [2pq] /(p_1 - p_2)^2 = [1.96 + 1.28]^2 \, [2(.5)(.5)] / (.2)^2 = 131$$

[Where p is the estimate sample proportion, and can be set at .5 if we make no assumption about the baseline and final coverage rates (p x q is maximum for p = q =.5).]

Statistical software, using formula (5) would yield $N_1' = N_2' = 140$.

140 children would be needed in the 12-23 months old group of interest for the baseline and final survey, if the samples were drawn by a simple random procedure. The cluster design forces us to correct the sample size in order to maintain the level of precision.

$$N_{1c} = N_{2c} = 140 \times 1.8 = 252$$

If we simply increased the total sample size in order to achieve 232 children in the 12-23 months age group, by the same process as in the preceding example, we would need a sample size for baseline and final evaluation of:

$$N_{1t} = N_{2t} = 252 \times (100 / 45) = 560 \text{ children.}$$

Example 5: Assessing the significance of an observed change

Let us now consider a situation where the baseline and final samples were chosen as:

$N_1 = N_2 = 166$

Our coverage rate estimates are 40% and 60% respectively at baseline and final, and we would like to assess whether this increase reflects a true change in the population of intervention. This question is similar to asking what the significance of the observed change is.
Our best estimate of the difference between the two proportions is: (.6) – (.4) = .2
If we ignored the deff, we could construct a 95% confidence interval for the difference between the two proportions, with the following formula:

$$95\%\ CI\ for\ (p_1 - p_2) = (p_1 - p_2) \pm Z\alpha \times \sqrt{[[(p_1q_1) / N_1] + [(p_2q_2) / N_2]]} \qquad \textbf{(6)}$$

In this case, we would obtain:

95% CI for (p1 - p2) = 0.2 ± 1.96 x √ [[(.4)(.6)) / 166] + [(.6)(.4)) / 166]]
95% CI for (p1 - p2) = 0.2 ± 0.105
Lower 95% CI for (p1-p2) = 0.095
Upper 95% CI for (p1-p2) = 0.305

- **The 95% CI (0.059 to 0.341) does not include zero, so we are 95% confident that a true increase of coverage rate has taken place. Our best estimate for this increase is 20%, and the 95% confidence interval is 9.5% to 30.5%.**

NOTES:

(a) Alternatively, a Z-test can be conducted to test whether the two proportions are equal to one another. This is equivalent to constructing a 95% CI and observing whether it includes zero or not.
$Z = (p_1 - p_2) / \sqrt{[[(pq) / N_1] + [(pq) / N_2]]}$
with $p = (N_1\ p_1 + N_2\ p_2) / (N_1 + N_2)$
Z can then be compared to a critical Z (e.g. 1.960 for a 5% significance level with one degree of freedom), which can be found in statistical tables.

(b) A more precise formula for the 95% CI is actually:
$95\%\ CI\ for\ (p_1 - p_2) = (p_1 - p_2) \pm Z_\alpha \times \sqrt{[[(pq) / N_1] + [(pq) / N_2]]}$
with $p = (N_1\ p_1 + N_2\ p_2) / (N_1 + N_2)$.

In reality, our sample did not come from a direct sampling method, and using the appropriate statistical software, the confidence interval would be corrected by a factor of the deff., that we will simply call C for this illustration.

True $95\%\ CI\ for\ (p_1 - p_2) = (p_1 - p_2) \pm Z_\alpha \times \sqrt{[[\mathbf{C} \times (p_1q_1) / N_1] + [\mathbf{C} \times (p_2q_2) / N_2]]}$

Depending on the deff, we would obtain a possibly much large CI, such as:

true 95% CI for (p1 - p2) ≈ 0.2 ± 0.205
Lower 95% CI for (p1-p2) ≈ -0.005
Upper 95% CI for (p1-p2) ≈ 0.405

The true 95% CI (-0.005 to 0.405) includes zero so, although our best estimate for the difference of coverage between the two phases is 20%, we cannot conclude that it is statistically significantly different from 0.

ANNEX 4: LQAS ALPHA AND BETA ERRORS (N=19)

Decision rules for an LQA sample of 19 in which alpha and beta errors for 30 percentage point differences in coverage are shown

	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
0%	1 0.014 0.000	2 0.031 0.000	3 0.046 0.000	3 0.017 0.000												
5%		3 0.111 0.067	3 0.046 0.067	4 0.059 0.013	4 0.023 0.013											
10%				4 0.059 0.115	5 0.070 0.035	5 0.028 0.035	6 0.032 0.009									
15%					5 0.070 0.144	6 0.078 0.054	6 0.032 0.054	7 0.034 0.016								
20%						7 0.173 0.068	7 0.084 0.068	7 0.034 0.068	8 0.035 0.023							
25%							8 0.180 0.077	8 0.087 0.077	8 0.035 0.077	9 0.035 0.029						
30%							8 0.180 0.182	9 0.184 0.084	9 0.088 0.084	9 0.035 0.084	10 0.033 0.033					
35%								9 0.184 0.185	10 0.186 0.087	10 0.087 0.087	10 0.033 0.087	11 0.029 0.035				
40%									10 0.186 0.186	11 0.185 0.088	11 0.084 0.088	12 0.077 0.035	12 0.023 0.035			
45%										11 0.185 0.184	11 0.084 0.184	12 0.077 0.087	13 0.068 0.034	13 0.016 0.034		
50%											12 0.182 0.180	12 0.077 0.180	13 0.068 0.084	14 0.054 0.032	14 0.009 0.032	
55%												13 0.175 0.173	14 0.163 0.078	14 0.054 0.078	15 0.035 0.028	16 0.013 0.008
60%													14 0.163 0.163	15 0.144 0.070	15 0.035 0.070	16 0.013 0.023
65%														15 0.144 0.150	16 0.115 0.059	16 0.013 0.059
70%															16 0.115 0.133	17 0.067 0.046
75%																17 0.067 0.111

How to use this table: Let us use the example of 80% coverage. Let's say you have a set a coverage target of 80% for a given indicator in your program. Go across the top row of the column and find 80%. Then go straight down until you find the place where it intersects with 50% (aka a coverage level 30% lower) on the left column. Here you will find 3 numbers in two rows. In the top row you will see the number 13, which is your decision rule in a sample size of 19 in an SA. This means that if you find 13 or more respondents in your survey who respond correctly to this survey you will classify this SA as having reached the target (although, as we know, there is a chance that it has fallen short). Now let's look at the second row where there are two numbers. The first is the alpha error (think of false negative) and is 0.068. This means that there is a 6.8 percent chance that **even though your SA has reached** 80%, you will find fewer than 13 respondents in your sample who respond correctly and you will **incorrectly classify the SA as not reaching the target** (when in fact it has). The second number in the row is the beta error (think of false positive) and is 0.084. This means that there is an 8.4 percent chance that you will find 13 or more respondents who respond correctly in an SA **even though its true coverage is 50% or below**. You will **wrongly classify this SA as having reached the target** of 80% when in fact it is far below. Each of the areas on this chart in which the intersection is in a bolded box show alpha and beta errors for differences in coverage of 30% (80/50, 75/45, 50/20, etc.). You will notice that for each of these cases the alpha and beta errors are below 10%. The other shaded areas demonstrate these errors for various coverage differences and show the corresponding alpha and beta errors in each case.

Similar tables for sample size 10-30 can be provided.