Outcome Monitoring and Evaluation Using LQAS

Participant Manual

USAID
FROM THE AMERICAN PEOPLE

US Agency for
International Development

APHIA II Evaluation
Made Possible By The Generous
Support Of The American People
Through USAID

MODULE OVERVIEW

Module 1: Program Planning Cycle (Role of Surveys and Linking Indicators to Plans)

The purpose of this module is to link surveys to program planning by helping participants see the link between (1) program objectives, (2) indicators used to assess progress towards them, (3) questions used to calculate indicators and, (4) the survey processes used to ask the questions.

Module 2: Random Sampling Background and Key Concepts and Issues

The purpose of this module is to demonstrate the importance of adhering to strict random sampling procedures (used in Lot Quality Assurance Sampling [LQAS]), to assure that participants understand what LQAS can and cannot tell them and to practice accurately describing LQAS results.

Module 3: Selecting Locations and Respondents

The purpose of this module is to practice how to randomly sample when a complete population list is not available using systematic sampling from a list and mapping within a community to select respondents.

Module 4: Obtaining, Tabulating, and Using Survey Results

The purpose of this module is to practice how to ask survey questions and then tabulate results and calculate indicators.

Module 5: Sampling Decisions, Logistics, and Reporting

The purpose of this module is to review sample size options when using LQAS and to consider key logistical and reporting issues related to surveys.

INTRODUCTORY MODULE

Outcome Monitoring and Evaluation Using LQAS

Introductions and Workshop Overview

Introductions, Norms, and Expectations

- Get to know a new friend and Write his/her Name, Organization, Zone on a name-card
- Take three notecards. Ask your friend three things s/he expects to learn from this workshop and write each of the expectations on its own notecard.
- Take one notecard and write on it three norms your friend would like to be observed during this workshop
- Introduce your new friend and give them the name-card to display
- Read out your friend's expectations and norms then hand in the respective notecards for posting on the wall (designated area)

Workshop Objectives

1. Review links between program planning, M&E, indicators and survey methods
2. Learn and practice random sampling
3. Learn and practice location and respondent selection using LQAS methodology
4. Learn and practice analysis and use of survey results
5. Understand the benefits and limitations to using LQAS as a methodology

Overview of the Schedule

Day	***Learning Activities***
Monday:	Introductions, program planning and M&E, introduction to random sampling
Tuesday:	Continuation of random sampling, understanding LQAS, selecting households and respondents
Wednesday:	Mobile solutions for surveys, parallel sampling, interview skills and field preparation for practice
Thursday:	Field practice, analyzing results
Friday:	Interpreting results, analyzing sample size options for LQAS, addressing logistical issues with implementing a survey, elements of a survey report, mobile solutions for conducting surveys

MODULE 1:

PROGRAM PLANNING CYCLE

Outcome Monitoring and Evaluation Using LQAS

Module 1:

Program Planning Cycle

(Role of Surveys and Linking Indicators to Plans)

Module 1: Objectives

1. **Articulate** how M&E fits within program planning and implementation processes

2. **Describe** the role of population-based surveys within program planning and M&E

3. **Select** or **develop** indicators to address knowledge, practice and coverage outputs

4. **Examine** examples of survey questions and questionnaire resources

Outcome Monitoring and Evaluation

Outcome Monitoring and Evaluation Workshop

Monitoring: Progress Tracking

- The *regular* collection and analysis of information
- Monitoring assists *timely decision-making*, ensures *accountability* and provide the basis for evaluation and learning.

  IFAD. (2002). *A guide for project M&E: Managing for impact in rural development.*
  (emphasis mine)

- Monitoring *assesses progress* against set objectives/outputs, *supervises* implementation and *assesses the effectiveness* of implementation strategies.

  The Applied Nutrition Programme, University of Nairobi School of Nutrition Science and Policy, Tufts University. (2000). *Monitoring and evaluation of nutrition and nutrition-related programmes: A training manual for programme managers and implementors.*
  (emphasis mine)

Sometimes there is confusion about the distinction between monitoring and evaluation. Monitoring is used as a way to **regularly or routinely assess progress** of program implementation to enable managers to make changes and move forward to improve implementation.

Evaluation

- A *systematic* (and objective as possible) examination of a *planned*, ongoing or *completed* project. It aims to *answer specific* management *questions* and to *judge* the overall value of an endeavor and supply lessons to improve future actions, planning and decision-making.

  IFAD. (2002). *A guide for project M&E: Managing for impact in rural development.*
  (emphasis mine)

- The process of determining the *worth* or *significance* of an activity, policy or program.

  Kusek, J.Z. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. The World Bank.
  (emphasis mine)

Evaluation is slightly different from monitoring in that it focuses on judging the worth of a program based on its plans. Note that, while they are different, they are both part of a process of systematically learning—monitoring in an ongoing way and evaluation as more of a point-in-time assessment.

Outcome Monitoring and Evaluation Workshop

- **Process:** Focuses on *outputs* (number of trainings, services offered, brochures distributed, visits made).
- **Outcome:** Focuses on changes in *knowledge, practice,* and *service* coverage.
- **Impact:** Focuses on attributing changes seen in a population to the program.

In considering the definition of "evaluation" one often encounters different kinds of evaluations: process evaluations, outcome evaluations or impact evaluations. Impact evaluations, because they seek to attribute change to the program require a complex (and often costly) randomization process whereby subjects are randomly assigned either to receive the program outputs (counseling, education or services for example).

Here we are not interested in merely assessing processes in the sense of counting the numbers of activities or services offered. We are also not going to to help you attribute causality. Rather, we will be talking about assessing changes in knowledge, practice, or service coverage of interest to your program. Even though you may not be able to prove that your efforts brought about these changes, we operate under the assumption that our program contributes to change, and so we want to assess whether change is actually occurring.

Research

- Whether you are doing monitoring or evaluation, you rely on the tools of research.

- Research is the *systematic process* of *collecting, analyzing, and interpreting* information (data) in order to increase our understanding of the phenomenon about which we are interested or concerned. (W)e intentionally set out to enhance our understanding of a phenomenon and expect to communicate what we discover.

  Leedy, P.D., & Ormrod, J.E. (2005). *Practical research: Planning and design.* Upper Saddle River, NJ: Prentice Hall. (page 2—emphasis mine)

We will be talking a great deal about following rigorous approaches. These come from the science of survey methodology. The processes are systematic and involve collecting, analyzing, and interpreting data.

Output M&E: What Really Matters

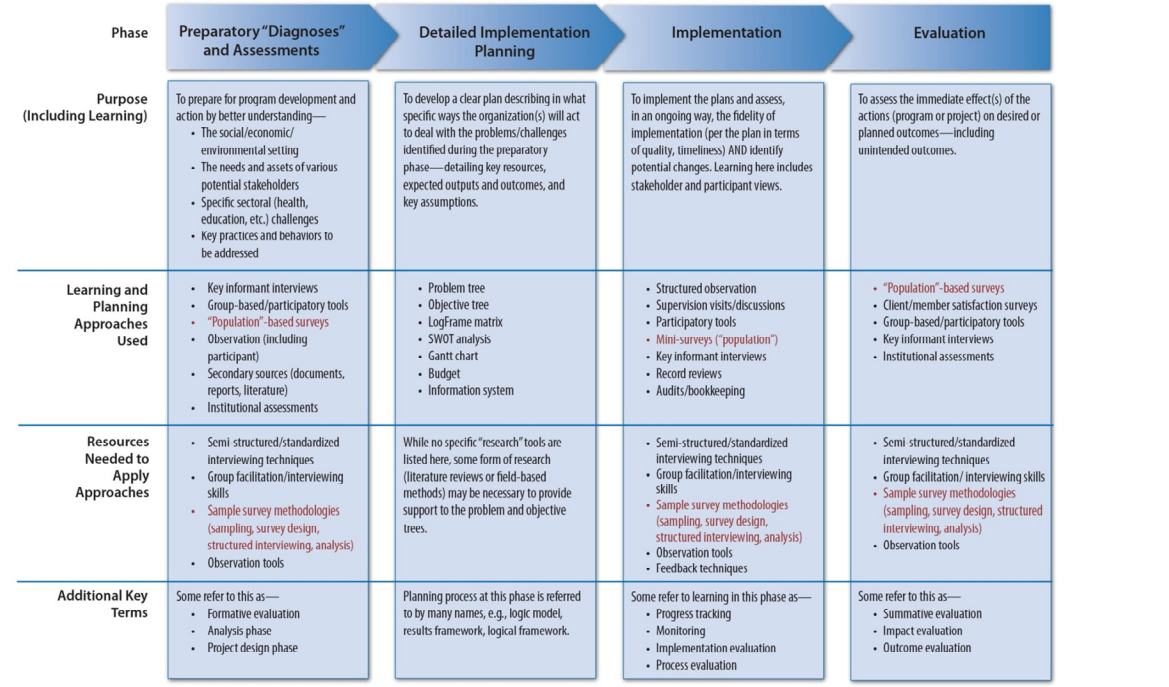
1. That you *rigorously* assemble evidence to…
   a. track your progress in an ongoing way (monitoring)
   b. assess the value of our work periodically (evaluation)
2. That you settle upon a *consistent and intelligent* method of assessing your output results, and then *tracking your trajectory* with *rigor.*

Collins, J. (2005). *Good to great and the social sectors.* Boulder, Colorado: Jim Collins. (emphasis in the original)

We will focus on rigor—especially in relation to sampling and making sure we are asking the right people the right questions. What we want is to have rigorous approaches so we can track progress—either in an ongoing way (monitoring) or to assess the value of what we are doing in a periodic way (evaluation).

As you develop your skills in sample survey methodology you will see that it allows you to have a consistent way to assess knowledge, practice and service coverage. Consistency allows you to compare results over time and by strictly following sampling, interviewing, tabulation and interpretation processes you achieve a rigorous approach to tracking your progress.

Program Planning and Implementation Cycle (Key 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, in an ongoing way, the fidelity of implementation (per the plan in terms of quality, 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") • 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 provide support to the problem and objective trees.	• Semi-structured/standardized interviewing techniques • Group facilitation/interviewing skills • Sample survey methodologies (sampling, survey design, structured interviewing, analysis) • Observation tools • Feedback techniques	• Semi-structured/standardized interviewing techniques • Group facilitation/ interviewing skills • Sample survey methodologies (sampling, survey design, structured interviewing, analysis) • Observation tools
Additional Key Terms	Some refer to this as— • Formative evaluation • Analysis phase • Project design phase	Planning process at this phase is referred to by many names, e.g., logic model, results framework, logical framework.	Some refer to learning in this phase as— • Progress tracking • Monitoring • Implementation evaluation • Process evaluation	Some refer to this as— • Summative evaluation • Impact evaluation • Outcome evaluation

This is a complex image, but it basically shows the entire program planning and implementation cycle—from planning to final evaluation. Note the points in red because they all refer to population-based surveys and use "sample survey methodology." We will focus on these methodologies, and we will define these terms.

The important thing to note in this slide is that the methods we will be discussing can be used before a program begins (baseline surveys) to set objectives and actually define the program. They can be used during implementation to monitor progress. They can also be used at certain times in the project (middle and end for example) to evaluate the project and its effects—its outcomes.

Please note also that there are other methods and tools that are critical to planning, monitoring and evaluation (ongoing learning) that we will not be discussing here. Sample survey methodologies are useful but not the only way to gain information about progress. Qualitative methods that rely on semi-structured (in-depth) interviewing are also very useful for example.

Do not worry if you do not know all the concepts on this slide. Here we are focusing only on "population-based" surveys and the methodologies used to accomplish them. We will go into these in detail.

Activity

1. Take a pause (10 minutes) to review the definitions shared in this session.

2. Put them into your own words—in a way you could explain them to someone else.

Population-Based Surveys

Population-Based Surveys

What are "population-based surveys?"

- Assess changes in the population by administering questions to people in the general population.
- The "population can be a subset of a larger population that the program/project is interested in, for example:
  - Young people age 15–24
  - Pregnant women
  - People living with HIV
  - Orphans
  - Others.
- These differ from service provider surveys, e.g., HFA, SPA (not the focus of this workshop).

To say we are doing a "population-based survey" is to say we want to have a way to assess (in our case) knowledge, practice and coverage in a population by surveying people. Other surveys focus on service providers (those who give HIV tests, those who provide prenatal or family planning services, or those who provide care to people living with HIV for example). While surveys of providers are important, that is not the focus of this workshop. We are interested in methods that allow us to find people in a population and ask them questions.

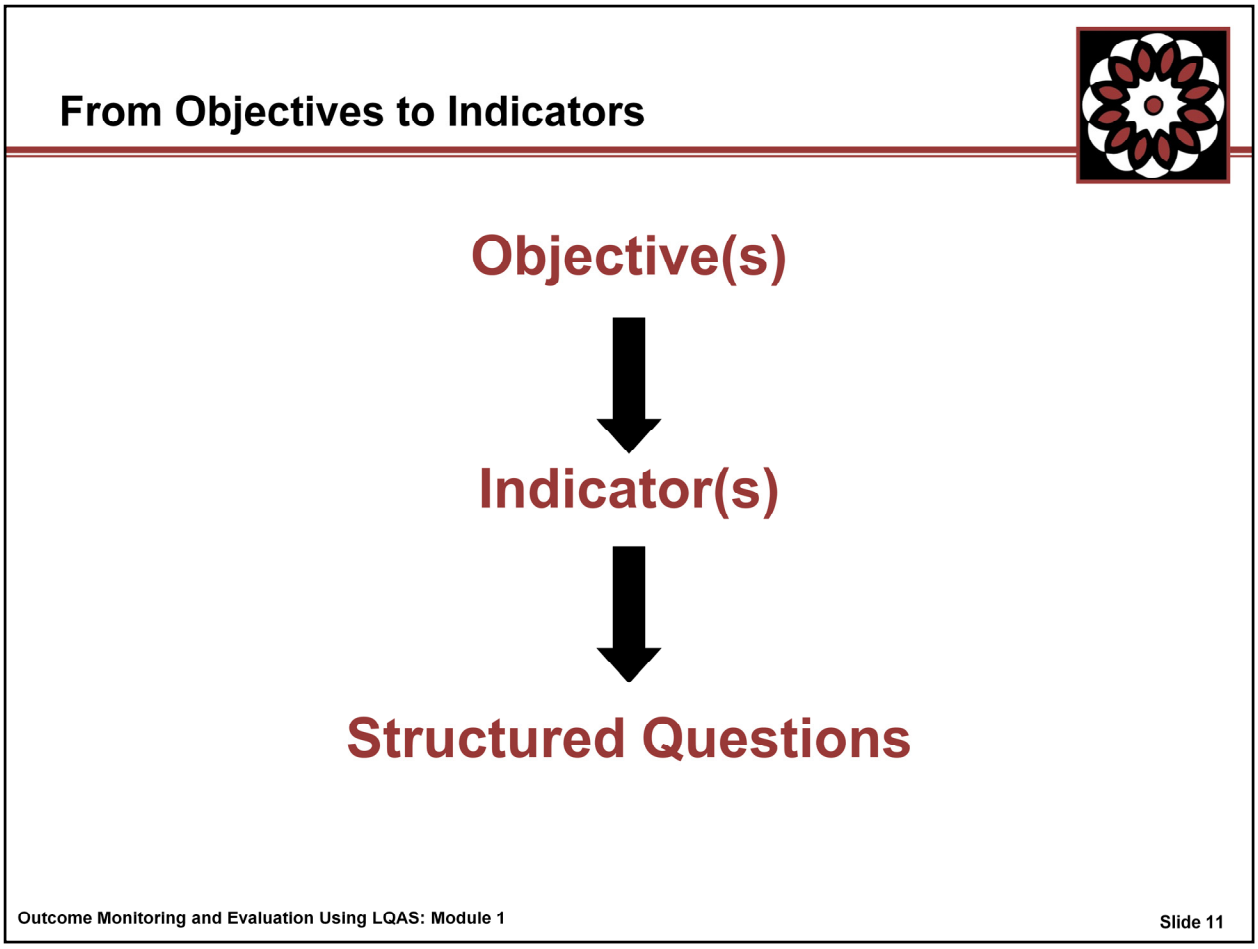
The idea of "population" can refer to the general population of people in an area or to subgroups we are particularly interested in working with—like young people or women of reproductive age. Thus, to say we are doing a population-based survey can mean that we are looking for a particular group within a population.

Activity

- Discuss the following in groups of three:
  - Your experiences in population surveys
  - What do those surveys help you to do?
  - What (more) would you have liked the surveys to help you to do?

- Individually,
  - Can you define in your own words what population-based surveys are?
  - What are your ideas on collecting and using data locally and for the entire program?
  - Write your responses to the above two issues on a notecard and post on the wall.

Linking Objectives, Indicators, and Questions



So far, we have been talking about how surveys—specifically, population- based surveys—fit into monitoring and evaluation activities. However, we have not talked yet about exactly what we are evaluating. This brings us to the issue of program planning and setting objectives for programs. It is not the purpose of this workshop to focus on setting objectives. We assume that your programs all have objectives related to HIV or reproductive health, and we want to help you develop the skills you need to do surveys that will enable you to determine whether you are accomplishing your objectives. However, objectives, as such, do not lend themselves to direct assessment. Objectives need to be fleshed out in a way that we can actually measure progress toward their achievement.

Typically, we state an objective and then we find one or more indicators that enable us to say whether there is evidence that we are meeting the objective. Indicators are not synonymous with objectives but are concrete measures that point us in a direction so we can see whether we are going toward an objective. Once we have an indicator, then we can develop one or more questions that will enable us to measure the indicator accurately. It is up to you to develop your program's objectives, but the good news is that, once you do so, there are many indicators that have been developed by others you can use to measure progress toward your objectives. Further, all these indicators are also linked to rigorously tested structured questions that you can take and adapt to your purposes.

From Objectives to Indicators

- Before writing an objective, you should have a broader "**Goal**," for example, "*reducing the transmission of HIV among young people or from mothers to their children*." This then leads to objectives.
- The objectives are basically clear statements of "the results" you aim to achieve so as to contribute to your broader goal.
- Indicators are like clues, signs, or markers that inform us on whether the program is achieving its results or objectives.

  a. An indicator measures one aspect of a program or project that is directly related to the program's results or objectives.
  b. The value of an indicator changes from baseline to the time of the evaluation.
  c. An indicator presents this change in a meaningful way, such as a percentage or number.
  d. Indicators should be measurable, precise, valid, and reliable.

Any program objective or set of objectives typically contributes to some broader goal that may be beyond the reach of the program. However, you should have a clear broader goal in mind and it should be clear how your objectives contribute to it. The goal may be beyond the lifespan of the program but your objectives should clearly contribute in some way to its attainment. As noted, objectives are the results you aim to achieve in your program.

Indicators, on the other hand, are the means we use to measure progress towards our indicators. The following slides give examples of different objectives that contribute to the same goal and possible indicators that could be used to measure each one.

PLEASE NOTE: there is no single indicator that must be used for a given objective. An objective may have several possible indicators that are used to assess it and, in fact, you may use several indicators to assess progress towards a given objective or result.

Notice how the indicators take the objectives and create very specific ways to measure them. For any given objective, there are possibly many indicators that could be used. These are just some common ones.

From Objectives to Indicators

- **Goal**

  Reduce the transmission of HIV among young people.

- **Objective**

  Increase the knowledge among young people of how HIV is transmitted in the next 5 years. (You could state the amount of change.)

- **Possible Indicator**

Knowledge

The percentage of respondents age 15–24 who, in response to a prompted question, say that people can protect themselves from contracting HIV by not having penetrative sex, by using condoms, or by having sex with only one faithful, uninfected partner.

Notice that objective here (and also those on the following slides) is linked to a broader goal—in this case reducing the transmission of HIV. Reducing the transmission of HIV may be difficult to measure, but the objective is something that a program can work to change. Programs can influence knowledge, practice or coverage defined by the objectives. The objectives here would, arguably, contribute to reducing the transmission of HIV—though we realize that knowledge alone will not lead to a reduction. Improved knowledge is a step, however.

From Indicator to Questions

Knowledge

The percent of respondents 15-24 who, in response to a prompted question, say that people can protect themselves from contracting HIV by having no penetrative sex, using condoms or having sex only with one faithful, uninfected partner.

Questions asked of 15-24 year olds in a survey:

Have you ever heard of an illness called AIDS?	Yes.....1 No.....2
Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?	Yes.....1 No.....2 Don't Know.....8
Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?	Yes.....1 No.....2 Don't Know.....8
Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all	Yes.....1 No.....2 Don't Know.....8

From Objectives to Indicators

- **Goal**

  Reduce the transmission of HIV among young people.

- **Objective**

  Increase the use of condoms during sex among unmarried young people in the next 5 years.

- **Possible Indicator**

Practice/Behavior

The percentage of respondents age 15–24 who say that they used a condom the last time they had sex with a non-marital, non-cohabiting partner, of those who have had sex with such a partner in the last 12 months.

Notice that the goal here is the same as in the last slide but that the objective is different. Here we are talking about a behavior (practice) change that our program is focused on. See also that a possible indicator focuses on the behavior in question and fully articulates what we mean in the objective when we say "increase the use of condoms among unmarried young people."

Notice also that the objective has a time element—in the next 5 years. This is common for objectives. What we are saying if we set this as a program objective is that we expect that if we increase condom use among young people that HIV transmission will be reduced. Our indicator is one way (a standard way, in fact) of assessing the objective.

In contrast to the first objective and indicator, which were about knowledge change, this one focuses on behavior change.

From Indicator to Questions

Practice/Behavior

The percent of respondents 15-24 who say they used a condom the last time they had sex with a non-marital, non-cohabiting partner, of those who have had sex with such a partner in the last 12 months.

Questions asked of 15-24 year olds who are not married in a survey:

Now I would like to ask you some questions about your recent sexual activity. Let me assure you again that your answers are completely confidential and will not be told to anyone. If we should come to any question that you don't want to answer, just let me know and we will go to the next question.			
When was the last time you had sexual intercourse?		Days _____ Weeks _____ Months _____ Years _____	
	Last Sexual Partner	**2nd to Last Sexual Partner**	**3rd to Last Sexual Partner**
When was the last time you had sexual intercourse with this person?		Days _____ Weeks _____ Months _____ Years _____	Days _____ Weeks _____ Months _____ Years _____
The last time you had sexual intercourse (with this second/third person), was a condom used?	Yes ..... 1 No (skip next question)..... 2	Yes ..... 1 No (skip next question)..... 2	Yes ..... 1 No (skip next question)..... 2
Was a condom used every time you had sexual intercourse with this person in the last 12 months?	Yes ..... 1 No 2	Yes ..... 1 No 2	Yes ..... 1 No 2

From Objectives to Indicators

- **Goal**

  Reduce the transmission of HIV from mothers to their children.

- **Objective**

  Increase the proportion of pregnant women who are counseled about and tested for HIV in the next 5 years.

- **Possible Indicator**

Coverage

The percentage of women who were counseled and offered voluntary HIV testing during antenatal care for their most recent pregnancy, accepted the offer of testing, and received their test results, of all women who were pregnant at any time in the (1 or) 2 years preceding the survey.

Here the broader goal to which the program wants to contribute is a bit different from the first two examples but still focuses on reducing the transmission of HIV. Here the focus is on the specific issue of reducing transmission between mothers and their children. The objective indicates that we think that if pregnant women are tested and counseled about their HIV status, this will contribute to the goal.

Notice, finally, that the indicator chosen here is about service coverage and is very specific—we want to measure the percent of pregnant women who were offered testing, accepted a test, and received the results.

This and the previous two slides demonstrate the link between a program's goal, its objective(s) and the indicators used to assess them. Each indicator makes it clear the population we would need to survey in order to calculate the indicator. This tells us whom we must sample. The next step is to see how each indicator requires one or more questions to be asked of the population of interest in order to calculate the indicator. You have examples of these in your handouts for the three indicators we have here. The main thing to remember is that, for the most part, you should not have to develop brand new questions to ask people. A great deal of work has gone into creating and testing questions and many references exist—including those in your resources—that you can draw on. Therefore, we will not spend more time on these issues. Rather, we will now talk about how to find people to survey—how to select them rigorously—and then how to tabulate and analyze results.

From Indicator to Questions

Coverage

The percent of women who were counseled and offered voluntary HIV testing during antenatal care for their most recent pregnancy, accepted an offer of testing and received their test results, of all women who were pregnant at any time in the (one or) two years preceding the survey.

Questions asked of mothers of children less than (1) 2 years of age in a survey:

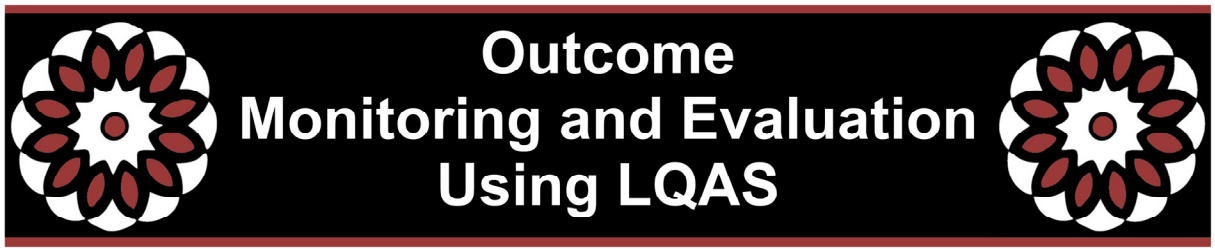
Did you see anyone for antenatal care before the birth of (NAME OF CHILD)?	Yes ..... 1 No 2
During any of the antenatal visits for that birth, did anyone talk to you about: A. Babies getting the AIDS virus from their mother? B. Things that you can do to prevent getting the AIDS virus? C. Getting tested for the AIDS virus?	Yes / No / DK A. AIDS from Mother: 1 / 2 / 8 B. Things to do: 1 / 2 / 8 C. Tested for AIDS: 1 / 2 / 8
Were you offered a test for the AIDS virus as part of your antenatal care?	Yes ..... 1 No 2
I don't want to know the results, but were you tested for the AIDS virus as part of your antenatal care?	Yes ..... 1 No 2
I don't want to know the results, but did you get the results of the test?	Yes ..... 1 No 2

Activity

- Get together with colleagues from your organization and do the following:
  - Look at the objectives of your program.
  - Examine from the list of indicators (**provided**) those that can best be used to measure progress toward your respective program objectives.
  - From the indicators suggested, select questions that can be asked to collect the necessary information from the population. (**provide questionnaires from where questions may be picked**)
- Share with the plenary how you have come down from objectives to the indicators, then to the questions.

MODULE 2:

RANDOM SAMPLING BACKGROUND, KEY CONCEPTS AND ISSUES



Module 2:

Random Sampling Background, Key Concepts and Issues

Module 2: Objectives

1. **Commit** to using random sampling in conducting population-based surveys
2. **Compare** and **contrast** two-stage LQAS and two-stage 30-cluster sampling
3. **Use** decision rules to assess LQAS results
4. **Analyze** what information LQAS can provide and its limitations
5. **Practice** describing LQAS results accurately

Random Sampling

Why Random Sample?

- Gives us results from the sample that reflect the real situation in the whole population (generalization/inference)
- Allows you to use the “few” to describe the “whole”
- Saves time
- Saves money
- Gives basis for use of statistical methods, e.g., calculating precision and confidence intervals

The reason sampling saves time and money is because we use a part of a population to draw inferences (conclusions) about the entire population rather than try to question the whole population. We can do this ONLY if we select our sample in a random way. We will talk more about this more but the key is to find ways to assure that randomness is assured at each stage in the process. Then, and only then, can we use sampling to:

-Use a relatively small number of people who are part of a population to make statements about the whole population.

-Save time

-Save money

Activity

1. Get into five groups of six people.
2. Your group will be provided two bags (A and B) each containing red and blue marbles.
3. Imagine each bag is a program area, and all the marbles in the bag represents all young people age 15-24 in that program area.
4. Each red marble in the bag denotes a young person who DOES NOT KNOW three ways of preventing HIV infection.
5. Each blue marble in the bag denotes a young person WHO KNOWS three ways of preventing HIV infection.
6. How can we find out what percentage of youth age 15-24 know at least three ways to prevent HIV infection in the program areas A and B respectively (i.e., the % of blue marbles in each bag)?

Activity—Continued

- To save time (and money) let us find out by randomly taking a sample of 30 marbles from the respective bags:
  1. Take bag A.
  2. Shake to mix the marbles.
  3. Close your eyes and pick 30 marbles.
  4. Count the number of blue and red marbles.
  5. Write the numbers on the flipchart.
  6. Are the young people who know how to prevent HIV infection "the majority," "just a few," or "somewhere in between"?
  7. Now empty the bag and find out the total number of blue and red marbles that were in the bag. Was your answer to the previous question a true reflection of the true situation in the bag?
- Repeat the above process with bag B.
- *What would the results have been if we used non-random sampling?*

Why Random Sample? (Recap)

- Gives us results from the sample that reflect the real situation in the whole population (generalization/inference)
- Allows us to use the “few” to describe the “whole”
- Saves time
- Saves money
- Gives basis for use of statistical methods, e.g., calculating precision and confidence intervals

In order to be able to interpret the results of a sample, we must use random sampling approaches. Without random sampling, we cannot use any of the statistical methods--including calculating the precision of our estimates with confidence intervals (which we will talk about later). Without random approaches, we cannot generalize our results to the entire population (use a “few” to describe the whole).

Hopefully the reason is now clear--without a random process our sample may include only those who are alike in some way in which we cannot know. They may be wealthier, or of a different age, or different in some other way compared to the entire population. Random selection is the only way we can make the assumption that our sample “represents” the broader population of interest.

Activity

- *We would like to ask you to indicate your commitment to using random sampling in surveys by writing your name and signing on the chart provided.*

Introducing Sampling Concepts

Activity: Reviewing Sampling Definitions

1. You are provided with a pack of 15 cards on each of which is a term usually used in sampling.
2. Take 10 minutes to read through the cards and stack them into piles:
   - Terms that you know
   - Terms that are new to you
   - Terms that are rather unclear to you
3. For the next 10 minutes take turns (2 minutes each turn) to show your partner one side of the flash card and see if s/he can give a response close to what is on the side of the card facing you.
4. Are there terms that are still unclear?

Introduction to LQAS

What is LQAS? Definition

- LQAS is
  - Lot
  - Quality
  - Assurance
  - Sampling
- Based on limited number of observations
- Can distinguish lots meeting pre-set outcome target from those that do not
- Used for monitoring, informs decision making on corrective measures
- Can be used aggregately to gauge coverage and outcome
- Adapted for public health in mid-1980s

The main things to take from this definition are:

1. LQAS relies on a relatively small sample;

2. It is useful for identifying program "sub-areas" (what we call supervision areas) that are failing to reach objective targets (or are performing at much lower levels than the program area as a whole), so we can focus attention on correcting problems in such areas;

3. It can be used to estimate overall levels of program knowledge, practice and coverage by combining results from program "sub-areas" (supervision areas).

What is LQAS? Background

- Developed in the 1920s to ensure that industrial processes produced and released goods meeting pre-set quality standards to the markets (outcome).
- Takes a small random sample of a manufactured batch (lot) and tests the sampled items for quality.
- The sample size is statistically determined (binomial probability)
- If defective items in the sample exceed a predetermined number (decision rule), then the lot is rejected.
- The decision rule is determined based on desired production standards.
- The sample size and decision rule give the manager high probability of rejecting substandard lots, and of accepting lots that meet the quality standards.

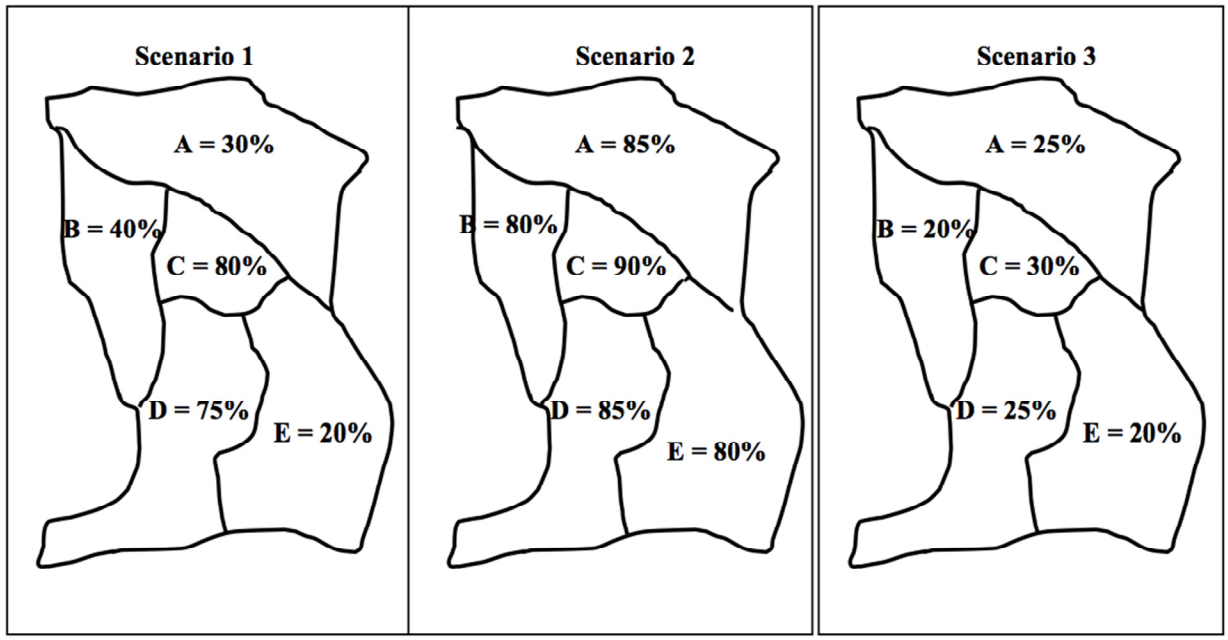
LQAS Definitions as Adapted for Public Health

Standard LQAS Theory	Public Health Programs
Production standard: % of items that must "pass" before the lot is accepted	Coverage Benchmark: % of persons to be covered by the service (i.e., received food or were vaccinated)
Production units: The machine or team that produced or assembled the lot	Supervision Areas: The CS or program **unit** responsible for delivering the service
Lot: Total number of items produced in given time by the production unit	Lot: Total number of persons in a given zone receiving service (food, vaccines, etc.)

LQAS Terminology as Applied in Programs

- Supervision Area (SA)
  - Catchment area or **program unit** to be assessed or monitored
- Coverage
  - Proportion with **desired outcome** in an indicator in a SA
- Coverage Benchmark
  - A preset **minimum** acceptable coverage level
- Average Coverage
  - Proportion showing desired outcome **across all SAs of the whole program area**
- Decision Rule
  - The **number** that corresponds to a specific coverage level for a given LQAS sample size

Uses of LQAS: Three Scenarios



Outcome Monitoring and Evaluation Using LQAS: Module 2 Slide 12

LQAS is most useful in distinguishing substandard supervision areas from others like in the first scenario. Let's assume that the program has set a target that in all supervision areas 70% of young people know three ways to prevent HIV transmission. In the first scenario supervision areas A, B and E are falling below this standard. LQAS is is less useful for distinguishing supervision when they are all performing at a high level or all at a low level as in Scenarios 2 and 3.

However, even in those cases it provides useful information to managers. If all areas are substandard vis-a-vis the program target (second scenario), then managers of various supervision areas can strategize about what to do.

If no areas are substandard (third scenario), then the program can continue without needing to worry about any particular attention paid to substandard performance. We use LQAS so that each supervision area within an entire program area will have some information to make decisions about whether they are substandard and in need of help. It is not meant as a judgment but to better help program managers decide where to target scarce resources.

30-Cluster Sampling and LQAS

Cluster Sampling

- Involves randomly selecting an interview location and sampling several respondents in it.
- In 30/10 cluster samples, you choose 30 interview sites and sample 10 respondents in each.
- You only have to go to 30 communities in your program area, and you can get results that tell you something about your entire program area.
- You do NOT get supervision-level information in 30-cluster sampling.

LQAS

- LQAS involves randomly sampling 19 interview locations in every supervision area where you have a program.
- You only interview one eligible respondent at each selected interview location.
- You can combine LQA samples to get information for the entire program area.
- LQAS helps managers and teams by giving them information to make decisions about JUST their area (their supervision-level information).

The reason for including this information and comparison is that 30-cluster sampling is commonly used in health programs. Cluster sampling was developed to enable programs to estimate overall levels of knowledge, practice or service coverage and had the advantage of limiting the number of locations to which surveyors need to go—thus reducing cost and time to conduct the surveys.

While LQAS is designed to provide some information at each supervision area—specifically, whether a given supervision area is failing to reach targets and in need of attention—30-cluster sampling only permits estimates of knowledge, practice or coverage levels for the entire program area.

LQAS also allows for estimates of knowledge, practice or coverage levels for the entire program area. This is what is meant by “you can combine LQA samples to get information for the entire program area. We will see how this works later.

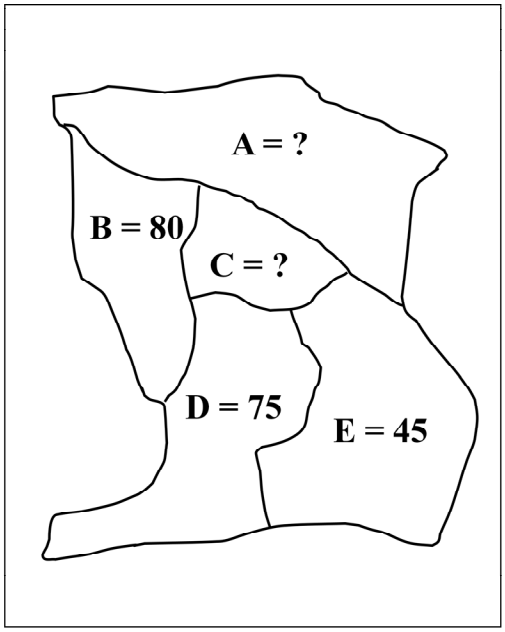
Because LQAS requires 19 interviews in each supervision area it is quite possible that, for a given program, it will require surveyors to go to more locations than required by 30-cluster sampling. This increases the cost of surveying for the entire program. However, each supervision area is required to go to only 19 locations AND they have some information they can use within their supervision area.

LQAS in Practice

NGO Program Area

Supervision Areas: A–E

Indicator: Percentage of young people (age 15–24) who know three ways to prevent HIV transmission



Again, we are working with a fictitious NGO Program Area, the same one that was used in the previous session. As you can see, we do not know the coverage for two of the five supervision areas—A and C—for the indicator "percentage of young people (15-24) who know three ways to prevent HIV transmission." Typically we would not know the coverage for any of the areas before doing a sample--that is why we sample. Here, however, we are assuming we do know the knowledge level for supervision areas B, D, and E.

Because we want to make decisions about deploying our program resources, we will need to do a survey of these two areas to see whether they need special attention. Saying they need attention means that we seek evidence that they are performing in a substandard way in relation to a target we have set for knowledge level (we have not yet defined this).

We don't know the coverage in supervision areas A or C, so we will do a survey, randomly selecting and interviewing 19 young people from each of the two supervision areas. We will demonstrate that this is adequate for our purposes of identifying priority areas.

Activity: Is a Sample Size of 19 adequate?

1. Form two groups.
2. Each group is provided with **two bags, A and C.** Each bag has red and blue marbles. Red marbles represent youth who DO NOT KNOW three ways of preventing HIV infection, and the blue marbles represent those who KNOW.
3. For each group, take a **random** sample of **19 marbles** from **bag A.** Count the blue marbles in the sample and record the number in the chart provided.
4. Repeat the process till you have taken **5 random samples of 19.**
5. Copy the sample results from the other group so you have a total of 10 random samples.
6. Now count all the marbles in the bag. How many blue marbles were in the bag? What percentage of all the marbles in the bag were blue?
7. Now let each group repeat this exercise using **bag C.**

LQAS Sampling Results

Indicator: Percentage of young people (age 15–24) who know three ways to prevent HIV transmission

Supervision Areas: NGO Program Area							
A				C			
# Correct (black draughts)				# Correct (black draughts)			
Sample		Sample		Sample		Sample	
1		6		1		6	
2		7		2		7	
3		8		3		8	
4		9		4		9	
5		10		5		10	

Verify "knowledge" in the bag for SA A → Total black draughts in the bag / Total black and red draughts in the bag = ☐/☐ = ____%

Verify "knowledge" in the bag for SA C → Total black draughts in the bag / Total black and red draughts in the bag = ☐/☐ = ____%

Optimal LQAS Decision Rules for Sample Sizes of 12–30 and Coverage Benchmarks of 10%–95%

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 can not 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.

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

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

On this table the first column (far left) is the size of your sample. Sample sizes of 12-30 are displayed. We are using a sample size of 19, and will why below. The percentages across the top of the page represent targets a program might have for knowledge, practice, or coverage.

The numbers at the intersection of the target (column) and row (sample size) are called “decision rules”, and they help us make a decision about a given supervision area. The way to think about the decision rule number is that it is the minimum number of people who must know, practice, or receive coverage for us to conclude that the supervision does not need immediate attention. In the example we have been using, it would be the minimum number of young people (in our supervision area sample of 19) that we would need to find that knows three ways to prevent HIV transmission in order to conclude that the supervision area does not need immediate attention.

Let’s say that we want 50% of young people to know three ways to prevent HIV transmission. Looking at the table, we find the number 7 at the intersection of the column of 50% and the sample size row of 19. If our target is 50% and we find less than 7 (the minimum number), we would conclude that this supervision area needs attention because there is evidence that it is falling below this target. On the other hand, if we find in our sample of 19 at least 7 young people who know three ways to prevent HIV transmission, we will conclude that the supervision does not need immediate attention. Now let’s say that our target is 80% knowledge. If the true knowledge level in the population is 80%, we should get at least 13 young people in a sample of 19 who know 3 ways to prevent HIV transmission almost every time. If we find fewer than 13 who know 3 ways we would conclude that this supervision area is in need of attention.

The Statistics of LQAS (I)

- If the true percentage of knowledge in the population were **80%,** *we would get 13 or more in a sample of 19 more than 90% of the time.* (We would get less than 13 less than 10% of the time.)
- At the same time, if the true percentage of knowledge in the population were **50%,** *we would get 13 or more in a sample less than 10% of the time.*
- So if our target is **80% knowledge of three ways to prevent HIV transmission among young people,** and we take a sample of 19, we can draw one of two conclusions:
  - If we get **13 or more,** we conclude that the *SA does not need attention at this time.*
  - If we get **less than 13,** we conclude that the *SA needs immediate attention.*

It is important to note that we never know the TRUE percentage in a population. If we did we would not need to do the survey! However, knowing that if the true percentage were a given level we would expect to get certain results with a known degree of certainty in a sample, enables us to use the results of the sample to make statements about the population that generated them. However, because the sample size in LQAS is relatively small, we must be very cautious in making those statements.

The way LQAS is set up (as we are using it), we can only conclude one of two things: either a supervision area is not in need of attention or there is evidence that it is underperforming and does require attention.

The Statistics of LQAS (II)

- If the true percentage of knowledge in the population were **50%,** *we would get seven or more in a sample of 19 more than 90% of the time.* (We would get less than seven less than 10% of the time.)
- At the same time, if the true percentage of knowledge in the population were **20%,** *we would get seven or more in a sample less than 10% of the time.*
- So if our target is **50% knowledge of three ways to prevent HIV transmission among young people,** and we take a sample of 19, we can draw one of two conclusions:
  - If we get **seven or more,** we conclude that the *SA does not need attention at this time.*
  - If we get **less than seven,** we conclude that the *SA needs immediate attention.*

This is merely a second example, using a different percentage target, of the same point made in the last slide. Recall that the number 7 (of 19) is taken from the table in the slide before last and concerns a sample of 19 if our target percentage is 50%. The number 7 is our decision rule.

Notice that it is very rare for us to conclude that a supervision area that is 30 percentage points below the target is NOT underperforming and in need of immediate attention. It happens less than 10% of the time. This means that if the true level of knowledge in the supervision area is 20% we would find 7 or more people with appropriate knowledge in our sample of 19 less than 10% of the time. In the same way, if the true level of knowledge in the supervision area is 50% we would find less than 7 people with appropriate knowledge in our sample of 19 less than 10% of the time.

Simplifying Our Conclusions

- Our target is 80% knowledge. If, in our sample of an SA, we find 13 or more who know three ways to prevent HIV transmission, **we classify the SA as not requiring priority intervention at this time.** If, however, we find fewer than 13 who know three ways to prevent HIV transmission, **we classify the SA as substandard and requiring immediate intervention.**

OR

- Our target is 50% knowledge. If, in our sample of an SA, we find seven or more who know three ways to prevent HIV transmission, **we classify the SA as not requiring priority intervention at this time.** If, however, we find fewer than seven who know three ways to prevent HIV transmission, **we classify the SA as substandard and requiring immediate intervention.**

The way that the LQAS tables you have are set up, you always know that there is less than a 10% chance of misclassifying a supervision area that has reached a target as in need of an immediate intervention, and there is less than a 10% chance of misclassifying a supervision area that has fallen 30 percentage points below the target as not being in need of an immediate intervention.

We have used the idea of “intervention” here. This word can mean many things in a given context, but it basically means you need to find out why the supervision area is performing in a substandard way in relation to the target you have set and think about steps to take to correct the situation.

In talking to staff and partners about how to interpret results from LQAS we suggest you use a formulation like the ones presented here. This will avoid problems of misinterpretation or drawing conclusions that a sample of this size is not designed to help you to draw.

Summary: LQAS and Why the Sample Size of 19

- LQAS is designed to give managers a signal to take immediate corrective action in a SA in relation to meeting the target on a given indicator.
- The LQAS table is designed to detect SAs falling at least 30% below the target as requiring immediate corrective action.
- The signal requires:
  - A target
  - A sample size
  - A decision rule
- Once we have two of these three requirements, the third is obtained from the LQAS table.
- The sample size of 19 is usually used because it is the smallest sample size with less than 10% alpha and beta errors across all coverage targets.

Below we discuss the issue of supervision areas that fall between less than 30 percentage points below the target. As noted here (and discussed above), LQAS is set up to identify supervision areas that are underperforming so we can provide more attention to them. Because of its small sample sizes it is particularly good at picking up those that fall far below the target. This is as it should be: we neither want to misclassify those that are falling far below as not in need of help nor do we want to discourage those that are reaching the target by classifying them as in need of help. Managers need to be able to identify those that are underperforming.

Just to be clear, a target refers to the level of knowledge, practice or service coverage you set for your program (this defines how you will decide that you are meeting your objective). The target refers to the level for each indicator and probably will vary from indicator to indicator. The sample size we are recommending is 19 because the probability of misclassifying in the two ways we have discussed is kept under 10% in both cases (this is what is meant by alpha and beta errors). The decision rule is taken from the table once you have decided on your target and sample size.

Uses and Limits of LQAS

What a LQAS Random Sample of 19 Can and Cannot Tell Us

1. The main use of LQAS is to determine if there are SAs within our program that are in need of immediate attention. **Thus, a sample of 19 helps us prioritize among SAs when there are large differences between them.** Specifically—
   - It accurately classifies substandard supervision areas as in need of a priority intervention.
   - It accurately classifies supervision areas that are *not* in need of a priority intervention as not needing one.
   - It helps us set priorities among different knowledge, practice, and coverage indicators within an SA.
2. In contrast, **a sample of 19 cannot help us to prioritize among supervision areas when there are small differences between them.** This does not mean LQAS is not useful, however, because we can use it to see if all areas are underperforming and/or to identify individual knowledge, practice, and coverage indicators that are underperforming.
3. We **cannot use a sample of 19** to calculate exact knowledge, practice, or coverage **for a single supervision area.** However, we can **combine individual samples of 19** to calculate knowledge, practice, and coverage percentages **for an entire program area.**

This slide is pretty self-explanatory but it does show the broader range of things that a sample of 19 in several supervision areas within a single program area can help us to do. We will come back to point three and discuss how this works below. For now let us make one comment about the third bullet point under the first item.

Up to now, we have been using an example of comparing a single indicator across many supervision areas. However, in your program you are likely to measure several indicators in each supervision area because your program is likely to have several objectives requiring several indicators. We have done it this way to keep things simple, but another use of LQAS is to assess, within a single supervision area, whether a given indicator requires immediate attention. We will provide an example of this later, but suppose you are assessing five indicators and have targets for each one. If for one or more of them you find that those with the appropriate knowledge, practice, or coverage falls below the decision rule for that target, you can prioritize that or those indicators (and program activities meant to help achieve them) for immediate action.

Why Use a Random Sample of 19?

A sample of 19 provides an acceptable level of error in two ways:

1. Less than 10% of the time, we will misclassify an SA that does **not** need immediate attention as needing a priority intervention.
2. Less than 10% of the time, we will misclassify and SA that **does** need immediate attention as not needing a priority intervention. *(Recall that this means that we will misclassify an SA that is 30 percentage points below the target as* not *needing a priority intervention less than 10% of the time.)*

Samples larger than 19 have practically the same level of accuracy as a sample size of 19. Thus, larger samples in a single SA (except for *much* larger ones) do not result in more accurate classification, and they cost more.

This slide merely summarizes much of what we have been saying. The key additional point here is that increasing the sample does not help one achieve much greater accuracy (in terms of the two types of misclassification we have been discussing). One would need to increase the sample size a great deal to achieve much greater accuracy but that inevitably increases the costs of doing the survey both in time and money.

Activity: Review LQAS and Sample Size 19

1. Discuss these two slides with a colleague.
2. Note down any points that are confusing or need to be clarified.
3. Share the points noted down with the rest.

Limits of LQAS

Let's say we want *all* SAs to achieve the result that at least 50% of all young people age 15–24 know three ways to prevent HIV transmission. If we take a sample of 19, what is the probability of misclassifying an SA as in need of a priority intervention (using a decision rule of 7) or of *not* in need of a priority intervention?

True Population Proportion who Know 3 Ways to Prevent HIV Transmission in the SA	Probability of classifying the SA as not needing an intervention (based on n=19 with decision rule of 7 or more who know 3 ways)	Probability of classifying the SA as needing an intervention (based on n=19 with decision rule of 7 or more who know 3 ways)
15%	2%	98%
20%	7%	93%
25%	18%	82%
30%	33%	67%
35%	52%	48%
40%	69%	31%
45%	83%	17%
50%	92%	8%
55%	97%	3%
60%	99%	1%
70%	100%	0%

We have discussed how LQAS rarely misclassifies supervision areas that have reached a target as "in need of attention." LQAS also rarely misclassifies supervision areas that are far below (30 percentage points) as not in need of attention. We now discuss those supervision areas that are below—but not too far below—the target. This slide shows the probability of misclassifying a supervision area if our target is 50% knowledge of three ways to prevent HIV transmission among young people (our decision rule is 7). Look first at the row for "true population proportion" of 50%. Notice the last column that says, probability of classifying the SA as needing an intervention. What this means is that if the true percentage of knowledge among young people is 50%, we would wrongly classify it as needing an intervention in only 7% of samples. This is what we meant we said that the probability of this kind of error is less than 10%. Now let's look at the line for 20%. Look at the second column: "probability of classifying the SA as not needing an intervention." Again, if the true percentage of knowledge among young people is only 20%, we would wrongly classify this SA as NOT needing an intervention only 7% of the time. This is what we meant when we said that the probability of this kind of error is less than 10%. You can see that if your target is 50% the chance of misclassifying a supervision area that is truly at 70% (by saying they need an intervention) is zero and that misclassifying an SA that is truly at 15% (by saying they do not need an intervention) is very small—only 2%.

The problem comes in the zone between 25 and 45%. Here, the closer the true population is to 50%, the more likely you are to classify it as not needing an intervention. You might say that this is not a problem because if it is close to 50%, it is not a priority—needing an immediate intervention. In fact, as it gets closer to 50%, this may be true. What a manager needs, and what LQAS is good for, is to identify supervision areas that are falling far below the target.

Limits of LQAS

Now let's say we want *all* SAs to achieve the result that at least 80% of all young people age 15–24 know three ways to prevent HIV transmission. If we take a sample of 19, what is the probability of misclassifying an SA in need of a priority intervention (using a decision rule of 13) or of *not* in need of a priority intervention?

True Population Proportion who Know 3 Ways to Prevent HIV Transmission in the SA	Probability of classifying the SA as not needing an intervention (based on n=19 with decision rule of 13 or more who know 2 ways)	Probability of classifying the SA as needing an intervention (based on n=19 with decision rule of 13 or more who know 2 ways)
25%	0%	100%
30%	0.1%	99.9%
35%	0.3%	99.7%
40%	1%	99%
45%	3%	97%
50%	8%	92%
55%	17%	83%
60%	31%	69%
65%	48%	52%
70%	67%	33%
75%	83%	17%
80%	93%	7%
85%	98%	2%
90%	99.8%	0.2%
95%	100%	0%

This slide illustrates the same points as the previous slide but this time with a target of 80% (a sample of 19 and decision rule of 13). Here notice the probability of misclassifying supervision areas that have a true knowledge level of between 55 and 75 percent. Again the probability of misclassifying them as not needing an intervention is higher as you approach 80%. However, as noted, the probability of misclassifying a supervision area with a true knowledge level of 50% with a decision rule of 13 is less than 10%—it is 8% to be precise.

While we call these "limits" of LQAS we should recall that the main purpose for using LQAS is to identify supervision areas that are falling far below the target so you can focus on them to improve their knowledge level.

Cumulative Binomial Probabilities for n=19

n=19	Cumulative Probabilities for Values of p with d "correct" responses (Probability of finding correct response or more out of 19 if the true percent is X																		
	True Percent in the Population																		
	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
Correct Responses out of 19																			
19	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.014	0.046	0.135	0.377
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.010	0.031	0.083	0.198	0.420	0.755
17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.005	0.017	0.046	0.111	0.237	0.441	0.705	0.933
16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.008	0.023	0.059	0.133	0.263	0.455	0.684	0.885	0.987
15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.010	0.028	0.070	0.150	0.282	0.465	0.673	0.856	0.965	0.998
14	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.011	0.032	0.078	0.163	0.297	0.474	0.668	0.837	0.946	0.991	1.000
13	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.012	0.034	0.084	0.173	0.308	0.481	0.666	0.825	0.932	0.984	0.998	1.000
12	0.000	0.000	0.000	0.000	0.000	0.003	0.011	0.035	0.087	0.180	0.317	0.488	0.666	0.818	0.923	0.977	0.996	1.000	1.000
11	0.000	0.000	0.000	0.000	0.002	0.011	0.035	0.088	0.184	0.324	0.494	0.667	0.815	0.916	0.971	0.993	0.999	1.000	1.000
10	0.000	0.000	0.000	0.002	0.009	0.033	0.087	0.186	0.329	0.500	0.671	0.814	0.913	0.967	0.991	0.998	1.000	1.000	1.000
9	0.000	0.000	0.001	0.007	0.029	0.084	0.185	0.333	0.506	0.676	0.816	0.912	0.965	0.989	0.998	1.000	1.000	1.000	1.000
8	0.000	0.000	0.004	0.023	0.077	0.182	0.334	0.512	0.683	0.820	0.913	0.965	0.989	0.997	1.000	1.000	1.000	1.000	1.000
7	0.000	0.002	0.016	0.068	0.175	0.334	0.519	0.692	0.827	0.916	0.966	0.988	0.997	0.999	1.000	1.000	1.000	1.000	1.000
6	0.000	0.009	0.054	0.163	0.332	0.526	0.703	0.837	0.922	0.968	0.989	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000
5	0.002	0.035	0.144	0.327	0.535	0.718	0.850	0.930	0.972	0.990	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.013	0.115	0.316	0.545	0.737	0.867	0.941	0.977	0.992	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.067	0.295	0.559	0.763	0.889	0.954	0.983	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	0.245	0.580	0.802	0.917	0.969	0.990	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1	0.623	0.865	0.954	0.986	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Use this table to determine the probability of finding "d" correct responses or more out of 19 for a given true percentage X in a population. So, for example, the probability of finding 13 or more "correct" responses out of 19 if the true population proportion is 80% "correct" is 0.932. The probability of finding 13 or more correct responses out of 19 if the true population is 50% is 0.084.

This slide formalizes what we have seen on the previous slide and demonstrates that these probabilities are derived from the binomial distribution. It tells us the probability of finding a result as great or greater than what we have found if the true percentage in the population is X. So, if the true percentage "correct" in a population is 80%, the probability of finding 13 or more "correct" responses is 0.932 (or we will find 13 or more 93% of the time).

The colored/highlighted cells in this table show what we saw in the previous slides and remind us of the fact that it is highly unlikely that we would fail to call a "50% population" substandard if our target is 80% and we take a sample of 19.

As noted, we would find 13 or more only 8.4% of the time (probability of finding 13 or more is 0.084) in a sample of 19 if the true population percentage correct is 50%. This table shows all cumulative (d or greater) probabilities for all percentages for n=19.

Activity: Describing an LQAS Result

1. Pair up.
2. Discuss the case study given in turns (each person leading discussion on one SA).
3. Share with the plenary your discussion of the SA (get 2 volunteers then others fill in).

Describing an LQAS Result

You are a program manager of an HIV program that has the goal of increasing appropriate testing and counseling concerning HIV for pregnant women in a region of your country. You have divided your program into 4 supervision areas and are working with providers to improve counseling and testing for HIV.

An indicator you have decided to track is the percent of pregnant women who are offered counseling and testing during their pregnancy. To assess the indicator you have developed a survey that you randomly administer to women who gave birth in the past year. At the beginning of the program you administered the survey and found that over the entire program area (all SAs combined) 25% of pregnant women were offered counseling and testing for HIV. Based on this you set a target that at the end of the program in four years 65% of women in the program area would be offered counseling and testing. You also set the target that after 2 years 40% would be offered counseling and testing

You are now two years into your program and have sampled and interviewed 19 women in each SA. You will examine the results of only two of the SAs here.

- Your decision rule is _____.
- In SA 1 you found that 7 of 19 women you interviewed (after random selection) said they had been offered counseling and testing.
- In SA 2 you found that 4 of 19 women you interviewed (after random selection) said they had been offered counseling and testing.

In each case what do you conclude at this time?

In pairs each person should take one of the SAs and in as simple terms as possible—and using what we have already seen and discussed—describe your result and what it implies for your program in that SA right now.

Use this space to write notes on how to accurately describe the results of a random sample of 19.

MODULE 3:

SELECTING LOCATIONS AND RESPONDENTS

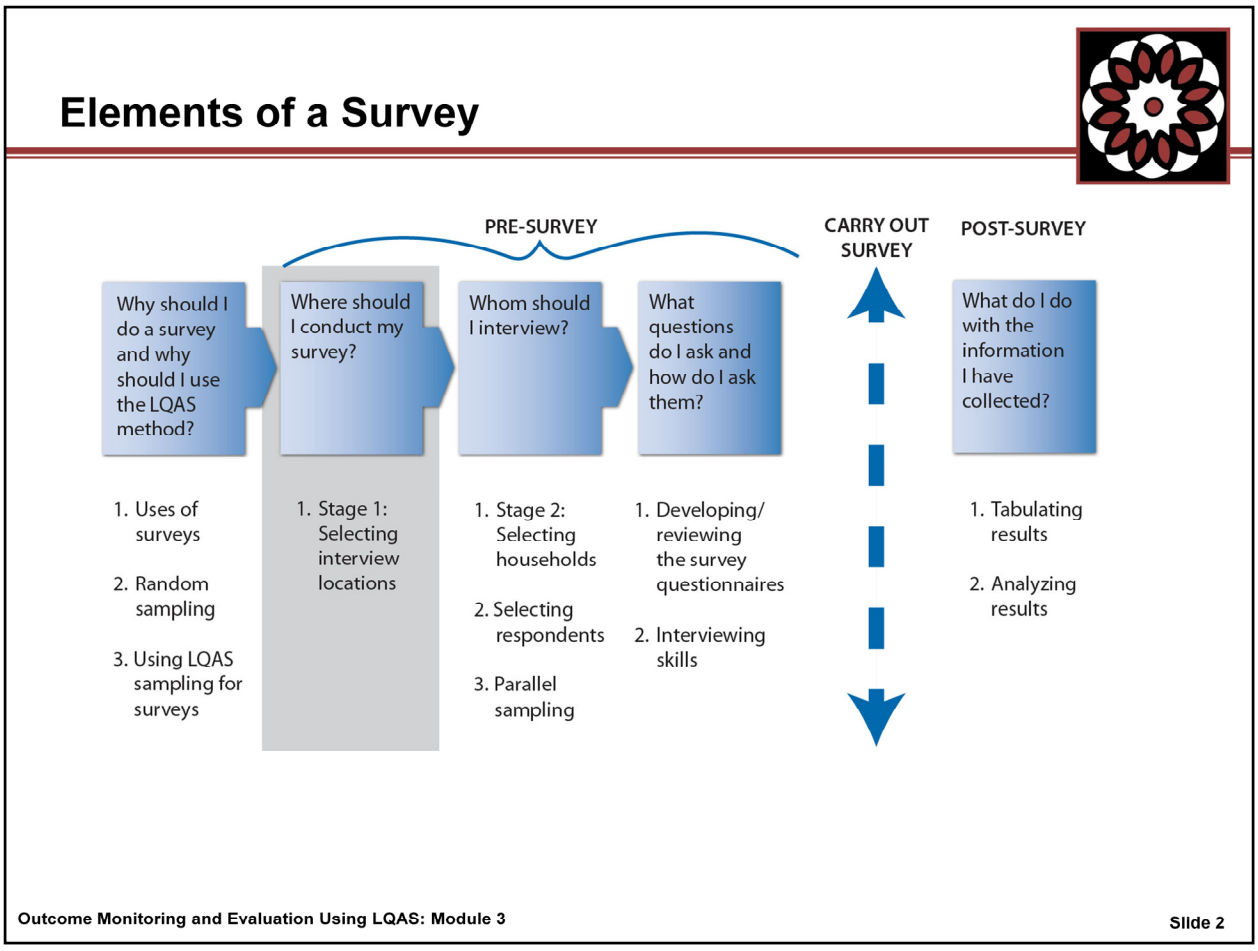
Outcome Monitoring and Evaluation Using LQAS

Module 3:
Selecting Locations and Respondents

Module 3: Objectives

1. **Select** interview locations for a single “Supervision Area” using a step by step process
2. **Assess** different scenarios for selecting “households” in a single “Supervision Area”
3. **Use** respondent identification rules to select respondents in different household composition scenarios
4. **Use** parallel sampling approach and rules for identifying different kinds of respondents in a single location

Identifying Interview Locations



This slide shows all the elements of a survey and up until now we have been focusing on the uses of surveys, random sampling and the specific sampling approach known as LQAS. In this section we discuss the first two elements in the "pre-survey" section: Where should I conduct my survey? and Whom should I interview? We are assuming that in your program you do not have a full list of the entire population from which your sample will be taken. This is most often the case for population-based surveys and so we must approach our work in several stages; first selecting communities or locations, then selecting households within the locations, and, finally, selecting respondents that fit the profile of people we want to survey. This final stage of selection depends on the indicator we want to measure and might be young people, women of reproductive age or adult men (for example).

We start by describing a step by step process for selecting interview locations.

Identifying Locations for Interviews
(with probability of selection proportional to size [PPS])

Step 1. **List communities and total population.**

Step 2. **Calculate the cumulative population.**

Step 3. **Calculate the sampling interval.**

Step 4. **Choose a random number.**

Step 5. **Beginning with the random number, use the sampling interval to identify communities for the 19 sets of interviews.**

We will walk through each step with a concrete example but pause here to note that to use this approach you will need both a list of all communities/locations in which you work as well as an estimate of the total population in each location. Even if you are going to sample a "sub-group" (young people or women for example) in this population we do not need the population of this subgroup since we assume they are represented in each community at about the same relative size.

Information on population size can come from a recent census in most cases unless there have been major population shifts in which case other estimates can be used such as those coming from other surveys or other sources that have needed to estimate the relative size of each community. We really need a sense of the relative size of each location in comparison to others as we shall see. We need to be able to select locations with their probability of selection proportional to their size (PPS) meaning that larger communities have a greater probability of being selected. Why sample with PPS? We do this because we want the distribution of the population we are interested in learning about (young people age 15-24 in the example we have been using) in the sample to mirror their distribution in the broader population.

The process we are using for the selection of locations is the same you would use if you were carrying out a 30-cluster survey. The difference between 30-cluster and LQAS is that in a 30-cluster sample, when we actually go to the interview sites we have selected in the first stage, we interview X individuals (say 10) in each cluster. By contrast, in LQAS, we interview only one person per selected site. Also, in LQAS, we are going to be selecting interview locations in a single supervision area. In 30-cluster sampling, we would select interview locations for the entire program area.

Worksheet: Calculating Cumulative Population

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548			
Santai	730			
Serina	686			
Mulrose	280			
Fanta	1,256			
Bagia	684			
Rostam	919			
Mt. Sil	1,374			
Livton	1,133			
Farry	544			
Tunis	193			
Pulau	375			
Sasarota	333			
Pingra	3,504			
Kanata	336			
Sirvish	2,115			
Balding	258			
Rescuut	678			
Krista	207			
Manalopa	1,162			
Garafa	408			
Spiltar	455			
Masraf	978			
Abrama	335			
Junagadh	541			
Singri	725			
Kalarata	355			
Ichimota	498			
Chaplar	347			
Sr. Kitt	186			
Nevis	1,346			
TOTAL	**23,489**			

Sampling Interval = _____

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = _____

This is a list of communities in a single supervision area, each shown with an estimated population. In the case of an urban area, we may have data for neighborhoods and, thus, could list them as separate communities. For survey purposes, we need to know only the total population of each community/neighborhood, not how many men vs. women, for example, or adults vs. children.

You should list communities in a random way—not by size or any other characteristic.

Begin by adding the population of the second community (Santai, 730) to that of the first (Pagai, 548) and writing the total (548 + 730 = 1,278) in the first blank space in the far right column ("Cumulative Population").

Repeat this process by adding the population of the third community (Serina, 686) to that of the combined population of Pagai and Santai (1,278) to get the new total (686 + 1278 = 1,964). Write it in the blank space in the far right column. Then do the same for the next community, Mulrose, adding its population (280) to the previous total (1,964) to get the new total: 280 + 1,964 = 2,244.

Worksheet: Calculating Cumulative Population, Choosing Interview Locations

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548	548		
Santai	730	1,278		
Serina	686	1,964		
Mulrose	280	2,244		
Fanta	1,256	3,500		
Bagia	684	4,184		
Rostam	919	5,103		
Mt. Sil	1,374	6,477		
Livton	1,133	7,610		
Farry	544	8,154		
Tunis	193	8,347		
Pulau	375	8,722		
Sasarota	333	9,055		
Pingra	3,504	12,559		
Kanata	336	12,895		
Sirvish	2,115	15,010		
Balding	258	15,268		
Rescuut	678	15,946		
Krista	207	16,153		
Manalopa	1,162	17,315		
Garafa	408	17,723		
Spiltar	455	18,178		
Masraf	978	19,156		
Abrama	335	19,491		
Junagadh	541	20,032		
Singri	725	20,757		
Kalarata	355	21,112		
Ichimota	498	21,610		
Chaplar	347	21,957		
Sr. Kitt	186	22,143		
Nevis	1346	23,489		
TOTAL	**23,489**			

Sampling Interval = _____

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = _____

This slide provides the entire cumulative population.

We will use the sampling interval to choose the interview locations by adding the interval amount to a random starting number (that we will select soon). You can see from this slide how to calculate the sampling interval. It is calculated by dividing the total population by the number of interview sites you need. In our case, that is 19 (per SA). In 30-cluster sampling, we would divide the population of the entire program area by 30 to get the sampling interval.

Worksheet: Calculating Sampling Interval

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548	548		
Santai	730	1,278		
Serina	686	1,964		
Mulrose	280	2,244		
Fanta	1,256	3,500		
Bagia	684	4,184		
Rostam	919	5,103		
Mt. Sil	1,374	6,477		
Livton	1,133	7,610		
Farry	544	8,154		
Tunis	193	8,347		
Pulau	375	8,722		
Sasarota	333	9,055		
Pingra	3,504	12,559		
Kanata	336	12,895		
Sirvish	2,115	15,010		
Balding	258	15,268		
Rescuut	678	15,946		
Krista	207	16,153		
Manalopa	1,162	17,315		
Garafa	408	17,723		
Spiltar	455	18,178		
Masraf	978	19,156		
Abrama	335	19,491		
Junagadh	541	20,032		
Singri	725	20,757		
Kalarata	355	21,112		
Ichimota	498	21,610		
Chaplar	347	21,957		
Sr. Kitt	186	22,143		
Nevis	1346	23,489		
TOTAL	**23,489**			

Sampling Interval = 1236.26

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = _____

This slide shows the calculated sampling interval. For now keep the two decimal points at the end.

We must now choose a random starting number. This number must be between 0 and the sampling interval. Choosing random numbers is a common task when conducting a survey using a random approach. In this particular instance, we are using a random number table to help us identify the first interview location. We will use a random number table that has 14 columns made up of rows of random numbers.

You can use any randomizing process but wish, but using a random number table is recommended. The number of the sampling interval fixed is 1236.26 . The random number has to be between 0 and the sampling interval 1,236. (The decimal point is not used in this step.) The highest possible number of digits in the random number this case is 4 the number of digits in the interval (1...2...3...6), so we will be selecting a number with four digits.

Random Number Table

87172	43062	39719	10020	32722	86545	86985	04962	54546	23138	62135	55870	97083	67875
28900	50851	30543	89185	16747	95104	49852	26467	58869	79053	06894	23975	34902	23587
86248	71156	55044	13045	33161	95604	57876	23367	10768	78193	60477	03[illegible]7	06498	48793
10531	51391	41884	69759	32741	70072	01902	96656	90584	59263	49[illegible]235		40055	20917
02481	90230	81978	39127	93335	74259	25856	52838	49847	69042	[illegible]	78159	40374	49658
23988	13019	78830	17069	58267	69796	94329	34050	20622	55[illegible]	40403	93790	77631	74261
37137	47689	82466	24243	10756	54009	44053	74870	28352	46389	38729	80349	50509	56465
38230	82039	34158	90149	82948	60686	27962	39306	53826	47852	76144	38812	76939	03119
98745	08288	19108	84791	58470	59415	45456	44839	8627				47169	95273
44653	58412	91751	14954	87949	81399	51105	29718	8278				42829	26308
88386	66621	16648	19217	52375	05417	26136	05952	7195				01377	47012
50660	58138	01695	69351	25445	20797	74079	60851	4763				58484	12506
36732	74234	84240	46924	62744	39238	78397	60869	2642				17293	45096
34187	78277	83678	34754	46616	45250	25291	04999	19717	60324	66915	03473	98329	82447
26095	98131	79362	39530	53870	87445	26277	90551	28604	39865	40686	05435	74511	69866
00067	74289	20706	74076	28206	36960	09231	82988	57062	35331	08212	68111	52199	05065
42104	26434	30953	15259	76676	63339	75664	23993	63538	34968	47655	44553	61982	13296
82580	46580	87292	23226	21865	60338	04115	33807	38395	98484	40387	69877	24910	13317
89266	14764	17681	68663	66030	12931	17372	35601	63805	55739	42705	30549	31697	33478
47100	92329	89435	69974	40783	52649	93444	41317	02749	19052	34647	92814	88046	34020
59566	26527	44706	85670	96223	36275	82013	82673	60955	62617	90214	24589	59715	57612
10946	24676	66513	56743	96911	89042	08263	70753	89045	39189	04306	06090	94515	17772
34013	69250	27977	84597	55192	65088	55739	35953	18533	39339	78037	32827	68269	69218
21606	11751	30073	71431	53569	27865	90215	34772	21779	11734	64313	49764	30816	56852
56620	92612	77157	90231	90144	29781	01683	52503	60080	73703	70080	80686	47379	33279
49238	90475	84356	87159	21222	40106	02671	52684	38514	68434	16407	58164	13341	48142
50738	21999	73539	51802	78179	27872	57937	29696	67783	29373	96563	74619	77099	17190
58761	21571	71692	19723	25088	10483	71430	47068	78378	80237	32113	09381	62931	29243
55335	71937	22025	33538	04648	74232	57839	62431	61835	04784	06732	34202	93497	72070
26515	31143	83795	78445	32869	31489	81587	90354	97672	70106	35008	37899	36246	97805
32625	36806	00082	26902	26250	28919	38054	49027	22209	42696	46980	17065	61288	30208
20311	96089	20141	30362	04980	32703	04202	91080	28660	89691	84660	73433	70169	11273
10941	73003	87930	85620	06956	38719	88711	61454	64076	13316	02203	54437	54306	78229
56982	46636	34070	30803	39095	80387	08971	25067	07377	70704	13629	68474	99229	05535
14661	10670	15811	00454	81124	46977	89983	48836	48182	17054	06344	24267	16686	21401
52760	78118	23277	29760	00099	97325	54762	43117	73199	19621	24599	11030	64809	35088
48874	20831	02286	73635	93771	54264	49801	22653	01524	84621	91023	64028	29278	15987
44817	77408	48447	25934	22912	43086	68126	92970	91833	26418	72454	97636	94593	07880
17896	79375	70883	70135	21589	51181	71969	32951	35036	17219	27357	96517	55307	84470
27166	22347	92146	92189	16301	15747	72837	59174	75024	39459	54910	95335	95013	47068
13665	30490	63583	73098	19976	03001	94645	40476	43617	85698	66512	42759	20973	98759
58644	73840	08103	97926	57340	63077	08114	10031	35668	21740	33787	44756	20527	65367
72570	36278	06602	56406	85679	85529	08576	50874	59706	01019	29980	56742	05356	04810
92041	68829	02163	59918	83041	71241	90678	79835	86324	13075	29913	99831	25688	53648
71240	74119	53090	23693	14007	90107	68804	54927	68964	26535	28184	21630	12362	67990

622

Outcome Monitoring and Evaluation Using LQAS: Module 3

Slide 7

First, notice that each “block” of random numbers has five digits. Since we need a four-digit number, you will have to decide whether you will take the first four or the last four on any given block of numbers. Once you decide that, you can use several methods for selecting a random number. For this exercise, we will use the final four numbers.

Normally you would close your eyes and hold a pencil in the air over the random number table. Then you would bring the pencil down on the table while keeping you eyes closed. The pencil should strike on or near a row of random numbers near one of the columns of numbers. If the final four numbers are in the interval 1-1236 that would be your random starting number.

Here we assume that the pencil came down on the column and row in the red box giving us a random starting number of 622.

Worksheet: Randomly Choosing the First Interview Location

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548	548		
Santai	730	1,278		
Serina	686	1,964		
Mulrose	280	2,244		
Fanta	1,256	3,500		
Bagia	684	4,184		
Rostam	919	5,103		
Mt. Sil	1,374	6,477		
Livton	1,133	7,610		
Farry	544	8,154		
Tunis	193	8,347		
Pulau	375	8,722		
Sasarota	333	9,055		
Pingra	3,504	12,559		
Kanata	336	12,895		
Sirvish	2,115	15,010		
Balding	258	15,268		
Rescuut	678	15,946		
Krista	207	16,153		
Manalopa	1,162	17,315		
Garafa	408	17,723		
Spiltar	455	18,178		
Masraf	978	19,156		
Abrama	335	19,491		
Junagadh	541	20,032		
Singri	725	20,757		
Kalarata	355	21,112		
Ichimota	498	21,610		
Chaplar	347	21,957		
Sr. Kitt	186	22,143		
Nevis	1346	23,489		
TOTAL	**23,489**			

Sampling Interval = 1236.26

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = 622

The fifth and final step in this process using a random number and sampling interval to identify locations of 19 interviews. You are now ready to combine the results of the third and fourth steps of this process to identify interview locations.

The location number of the first interview is the random number we selected from the random number table: 622.

The location number of the second interview is equal to the random number plus the sampling interval, in this case 622 + 1,236.26= 1,858.26 (for this step you use the decimal).

The location number of the third interview is equal to interview location number 2 plus the sampling interval, or 1,858.26 + 1,236.26 = 3,094.52 and so on.

The following slides show how this continues. Make sure you can do the calculations.

Worksheet: Identifying the Location of Each of the 19 Interviews in a Supervision Area

LQAS No.	Calculation	Interview Location
1.	Random Number = Location Number 1	Random Number **622**
2.	RN + Sampling Interval = Location Number 2	622+1236.26 **1858.26**
3.	Interview Location Number 2 + Sampling Interval	1858.26+1236.26 **3094.52**
4.	Interview Location Number 3 + Sampling Interval	Etc.
5.	Interview Location Number 4 + Sampling Interval	
6.	Interview Location Number 5 + Sampling Interval	
7.	Interview Location Number 6 + Sampling Interval	
8.	Interview Location Number 7 + Sampling Interval	
9.	Interview Location Number 8 + Sampling Interval	
10.	Interview Location Number 9 + Sampling Interval	
11.	Interview Location Number 10 + Sampling Interval	
12.	Interview Location Number 11 + Sampling Interval	
13.	Interview Location Number 12 + Sampling Interval	
14.	Interview Location Number 13 + Sampling Interval	
15.	Interview Location Number 14 + Sampling Interval	
16.	Interview Location Number 15 + Sampling Interval	
17.	Interview Location Number 16 + Sampling Interval	
18.	Interview Location Number 17 + Sampling Interval	
19.	Interview Location Number 18 + Sampling Interval	

Sampling Interval = 1236.26

Random Starting Number = 622

This slide shows you how to get started with the calculations.

Worksheet: Identifying the Location of Each of the 19 Interviews in a Supervision Area

LQAS No.	Calculation	Interview Location
1.	Random Number = Location Number 1	622.00
2.	RN + Sampling Interval = Location Number 2	1858.26
3.	Interview Location Number 2 + Sampling Interval	3094.52
4.	Interview Location Number 3 + Sampling Interval	4330.78
5.	Interview Location Number 4 + Sampling Interval	5567.04
6.	Interview Location Number 5 + Sampling Interval	6803.30
7.	Interview Location Number 6 + Sampling Interval	8039.56
8.	Interview Location Number 7 + Sampling Interval	9275.82
9.	Interview Location Number 8 + Sampling Interval	10512.08
10.	Interview Location Number 9 + Sampling Interval	11748.34
11.	Interview Location Number 10 + Sampling Interval	12984.60
12.	Interview Location Number 11 + Sampling Interval	14220.86
13.	Interview Location Number 12 + Sampling Interval	15457.12
14.	Interview Location Number 13 + Sampling Interval	16693.38
15.	Interview Location Number 14 + Sampling Interval	17929.64
16.	Interview Location Number 15 + Sampling Interval	19165.90
17.	Interview Location Number 16 + Sampling Interval	20402.16
18.	Interview Location Number 17 + Sampling Interval	21638.42
19.	Interview Location Number 18 + Sampling Interval	22874.68

Sampling Interval = 1236.26

Random Starting Number = 622

This worksheet merely shows all the calculations and the number interview. It is a calculation sheet only.

Worksheet: Identifying Communities with Chosen Interview Locations

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548	548		
Santai	730	1,278	622	
Serina	686	1,964	1858.26	
Mulrose	280	2,244		
Fanta	1,256	3,500	3094.52	
Bagia	684	4,184		
Rostam	919	5,103		
Mt. Sil	1,374	6,477		
Livton	1,133	7,610		
Farry	544	8,154		
Tunis	193	8,347		
Pulau	375	8,722		
Sasarota	333	9,055		
Pingra	3,504	12,559		
Kanata	336	12,895		
Sirvish	2,115	15,010		
Balding	258	15,268		
Rescuut	678	15,946		
Krista	207	16,153		
Manalopa	1,162	17,315		
Garafa	408	17,723		
Spiltar	455	18,178		
Masraf	978	19,156		
Abrama	335	19,491		
Junagadh	541	20,032		
Singri	725	20,757		
Kalarata	355	21,112		
Ichimota	498	21,610		
Chaplar	347	21,957		
Sr. Kitt	186	22,143		
Nevis	1346	23,489		
TOTAL	**23,489**			

Sampling Interval = 1236.26

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = 622

622.00
1858.26
3094.52
4330.78
5567.04
6803.30
8039.56
9275.82
10512.08
11748.34
12984.60
14220.86
15457.12
16693.38
17929.64
19165.90
20402.16
21638.42
22874.68

This slide merely takes the results from the previous worksheet and shows how to place them in the appropriate community. Please note that the number of the interview location must be "within" the interval implied by the cumulative population. So, for example, Pagai's interval is 1-548; Santai's interval is 549-1278; Serina's interval is 1279-1964; etc. The first number--622--is in the interval between 549-1278 and that is why Santai is the first interview location.

Worksheet: Identifying Communities with Chosen Interview Locations

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548	548		
Santai	730	1,278	622.00	
Serina	686	1,964	1858.26	
Mulrose	280	2,244		
Fanta	1,256	3,500	3094.52	
Bagia	684	4,184		
Rostam	919	5,103	4330.78	
Mt. Sil	1,374	6,477	5567.04	
Livton	1,133	7,610	6803.30	
Farry	544	8,154	8039.56	
Tunis	193	8,347		
Pulau	375	8,722		
Sasarota	333	9,055		
Pingra	3,504	12,559	9275.82,10512.08,11748.34	
Kanata	336	12,895		
Sirvish	2,115	15,010	12984.60,14220.86	
Balding	258	15,268		
Rescuut	678	15,946	15457.12	
Krista	207	16,153		
Manalopa	1,162	17,315	16693.38	
Garafa	408	17,723		
Spiltar	455	18,178	17929.64	
Masraf	978	19,156		
Abrama	335	19,491	19165.90	
Junagadh	541	20,032		
Singri	725	20,757	20402.16	
Kalarata	355	21,112		
Ichimota	498	21,610		
Chaplar	347	21,957	21638.42	
Sr. Kitt	186	22,143		
Nevis	1,346	23,489	22874.68	
TOTAL	**23,489**			

Sampling Interval = 1236.26

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = 622

622.00
1858.26
3094.52
4330.78
5567.04
6803.30
8039.56
9275.82
10512.08
11748.34
12984.60
14220.86
15457.12
16693.38
17929.64
19165.90
20402.16
21638.42
22874.68

Pingra will be the location of three interviews because of its relatively large total population. This is not a problem. When the interview team arrives in Pingra (and we will see that the same thing will happen in Sirvish), they can divide the community into natural neighborhoods with help from an informant. They can then randomly select three neighborhoods. Each neighborhood, in this approach, should have roughly the same population in it, and every neighborhood should be included. It may be possible to find census information about Pingra that breaks down the community into census blocks. If so, you can use that information and the same approach we used here to select neighborhoods with probability proportionate to size. This illustrates, again, the point that we are selecting locations with probability proportional to size. In this case it means that there are several interview locations in a single community. This may not always be the case depending on the size and number of communities in a supervision area.

Worksheet: Summarizing Number of Interview Locations per Community

Name of Community	Total Population	Cumulative Population	Interview Locations	Number of Interviews
Pagai	548	548		
Santai	730	1,278	622.00	1
Serina	686	1,964	1858.26	1
Mulrose	280	2,244		
Fanta	1,256	3,500	3094.52	1
Bagia	684	4,184		
Rostam	919	5,103	4330.78	1
Mt. Sil	1,374	6,477	5567.04	1
Livton	1,133	7,610	6803.30	1
Farry	544	8,154	8039.56	1
Tunis	193	8,347		
Pulau	375	8,722		
Sasarota	333	9,055		
Pingra	3,504	12,559	9275.82,10512.08,11748.34	3
Kanata	336	12,895		
Sirvish	2,115	15,010	12984.60,14220.86	2
Balding	258	15,268		
Rescuut	678	15,946	15457.12	1
Krista	207	16,153		
Manalopa	1,162	17,315	16693.38	1
Garafa	408	17,723		
Spiltar	455	18,178	17929.64	1
Masraf	978	19,156		
Abrama	335	19,491	19165.90	1
Junagadh	541	20,032		
Singri	725	20,757	20402.16	1
Kalarata	355	21,112		
Ichimota	498	21,610		
Chaplar	347	21,957	21638.42	1
Sr. Kitt	186	22,143		
Nevis	1,346	23,489	22874.68	1
TOTAL	**23,489**			**19**

Sampling Interval = 1236.26

(Sampling Interval = Cumulative Population/Number in Sample 23,489/19)

Random Starting Number = 622

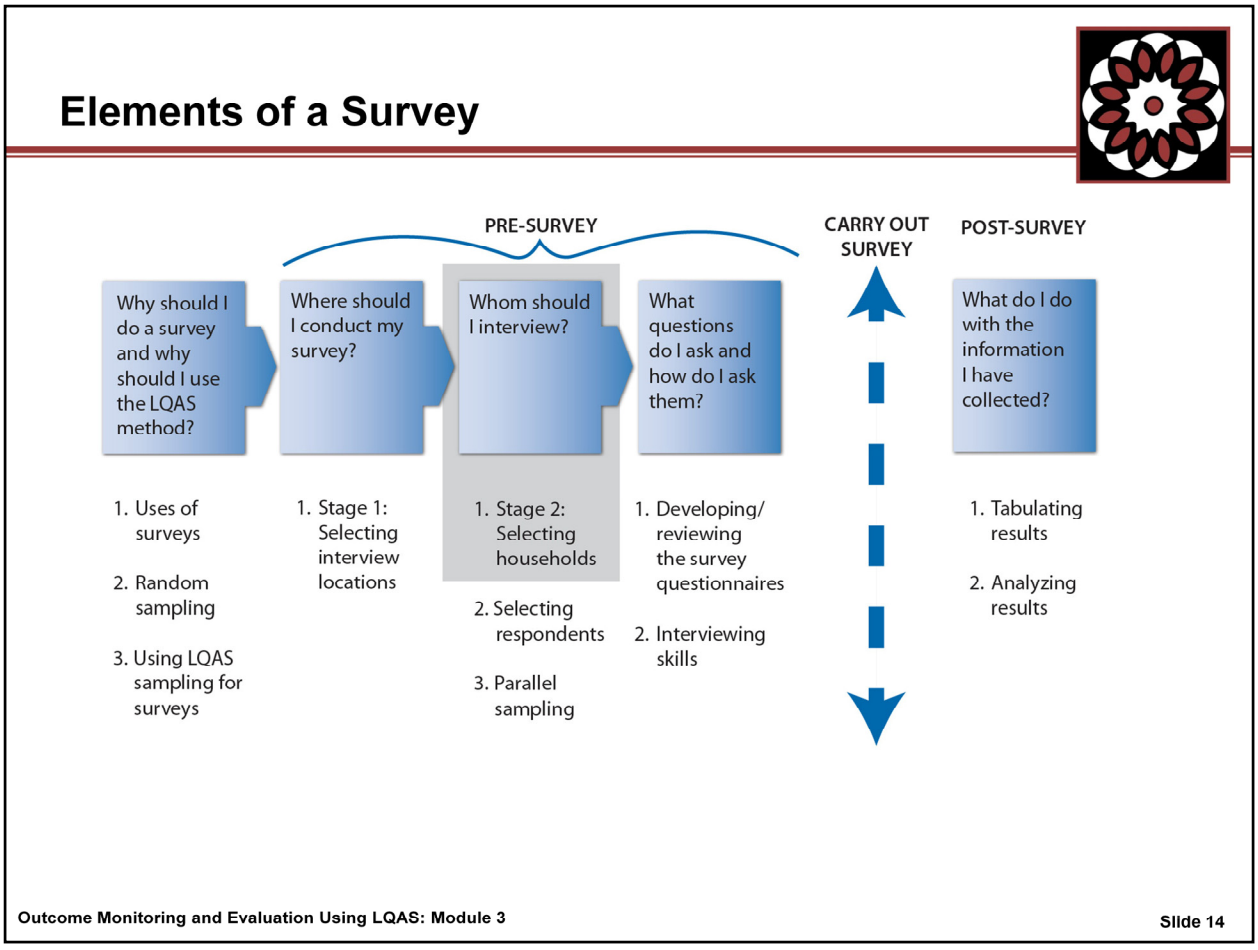
622.00
1858.26
3094.52
4330.78
5567.04
6803.30
8039.56
9275.82
10512.08
11748.34
12984.60
14220.86
15457.12
16693.38
17929.64
19165.90
20402.16
21638.42
22874.68

This slide shows the communities selected and the number of interview locations in each. In most cases there is one interview location per community but both Pingra and Sirvish have more than one.

Now that we have selected the interview locations for our 19 interviews in the supervision area (and we would do the same thing we just did for each one), we are ready to go to the locations and find the people we want to interview. The second stage of selection takes place in the locations we selected in the first stage, and in the next session, we will talk about how to do that. We will select households in the second stage.

One final note: This is the process you will go through for a two-stage sample, and you use this when you know where your population lives (all communities) but have no other way to easily identify them. If you have a program and you have a list of all participants—for example, if you are providing services to orphans and have all their names in a register—you can simply divide the list of names up by supervision area and randomly select 19 from each supervision area. In that case, you would not make a community list but make a numbered list of participants in each supervision area, select 19 numbers at random, and select the people from the list that correspond to the numbers you selected. So, if you had 150 orphans in supervision area 1, you would select 19 random numbers between 1 and 150, and find the names that correspond to those numbers on your numbered list. You can then go and interview these people. This is not multistage sampling but something called simple random sampling from a list.

Selecting Households & Respondents



We are assuming that we have identified the locations for the 19 interviews and have now gone to the first location. Our first task is to select a household at random, and this session will teach us how to do that. We selected interview locations randomly and we must continue to use random processes to select households and then individuals within them.

Please note again that you can use the process we are describing here to select the first household in an LQAS or in a 30-cluster sample. In a 30-cluster sample, you will need to have an approach to select subsequent households after the first one. We will describe options for doing this in a moment.

How to Assign Numbers to Households

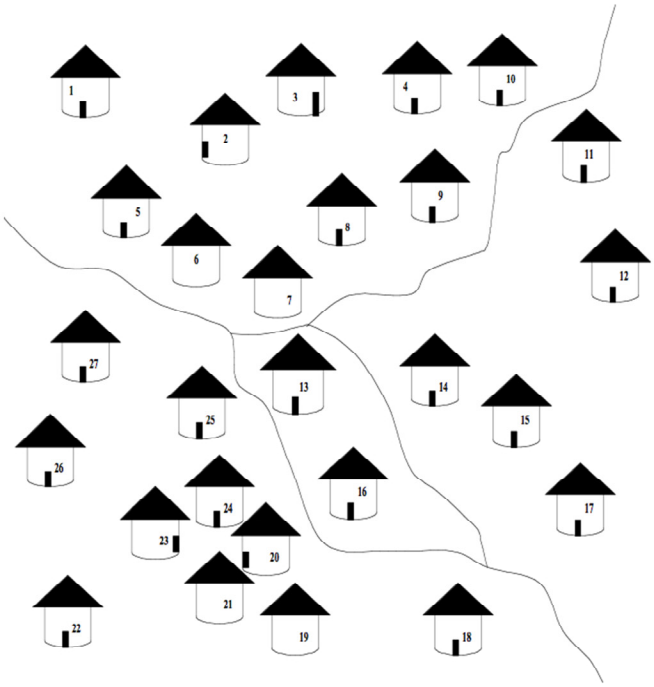
If:	Then:
A complete household* list is available (tax list, census, map)	Assign a number to each house. Select a random number between 1 and the highest number. Work is done!
If the community size is "about" 30 households or less	Make a household list or map with the location of each household (use assistance of a key informant from the community). Then assign a number to each house. Select a random number between 1 and the highest number. Work is done!
If the community size is more than "about" 30 households	Subdivide the community into 2–5 sections with about the same number of households in each section. Select one section at random. If the section has more houses than you can easily count, subdivide into 2–5 sections again and select one at random. Continue doing this until about 30 houses remain. Make a house list or map with the location of each household (use an assistant or key informant from your community). Then assign a number to each house. Select a random number between 1 and the highest number. Work is done!

* Household = Group of persons who share the same kitchen or hearth; a group of persons who eat from the same cooking pot.

We begin by numbering households in an interview location so we can randomly select one. There are three scenarios here, and we are going to spend time on the second and third because they are the most common in our work. If a community has a complete list of households, which is what we see on Line 1 of the slide, your task is easy. You can merely place a number by each household on the list, and then choose a random number between one and the highest-numbered household. We will talk about this in a moment.

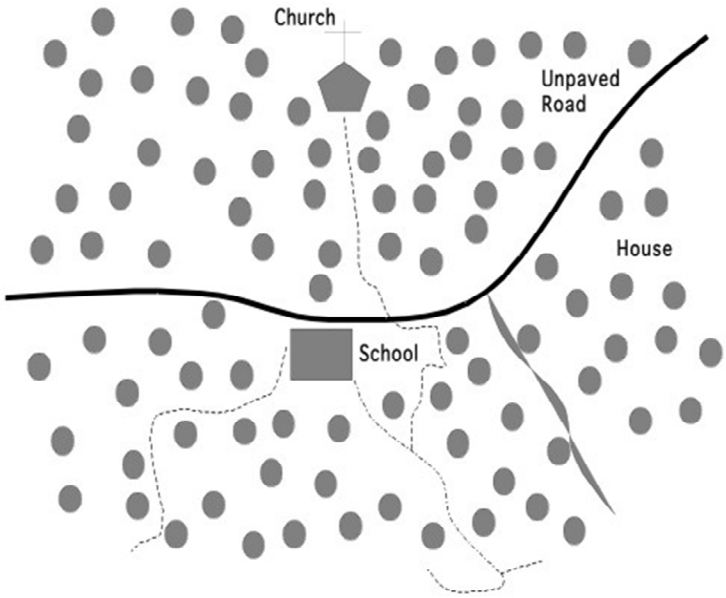
Having an updated list of all households in a community is rare, so you will probably need to create a simple map following the instructions on this slide for larger and smaller communities.

Situation 2: Household List Not Available—Size About 30



In this case (about 30 households in a community), the interviewer will have to draw a map of all the households in that location, with the help of an informant (that is, someone who lives in the community), and then assign numbers to the houses on the completed map. If a map is available, however, review it with the informant to make sure it is accurate and then assign numbers. Once the numbers are assigned select a random number from 1 to highest numbered household and go to that household.

Situation 3: Household List Not Available—More Than 30



In this case (more than about 30 households in a community), the interviewer will

1. Subdivide the community into two or more equal sections (you do not need to create a map of the entire interview location, but should be sure that subdivisions are roughly equal in size);

2. Select one of these sections at random;

3. If the selected area is still too large, subdivide it again into two or more equal sections, number each section, and select one section at random;

4. Continue until you have one small section with 15-25 households;

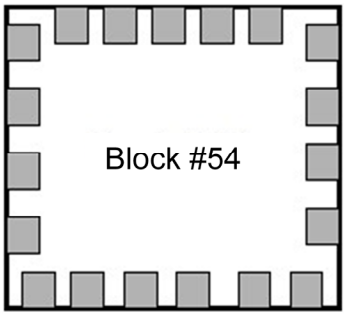
5. Draw a map of the section with the help of an informant; and

6. Number the households in this section on the map (you only need to count the houses in the selected section). Then you can randomly select a household by selecting a random number from 1 to highest numbered household.

Now that we have selected the first household, we need to think about what to do to select subsequent households. We might need to do this for several reasons. First, in an LQAS, the first household we select may not have the kind of respondent we are looking for. For example, to go back to the example we have been using, if we want to interview someone age 15-24 and our selected house does not have anyone of that age, we would need to select another household. We will discuss this below. Second, in a 30-cluster sample, we always plan to interview respondents in more than one household, so we will need a way to select the second and subsequent household. An accepted way to do this is to stand at the front door of the household we have just been in and select the household whose door is closest to the one where we are. This does not mean we can actually see the door; it may be on the other side of the household. It merely means that the door is the closest to the door of the household where you are.

Selecting the First House: Urban Blocks or Districts

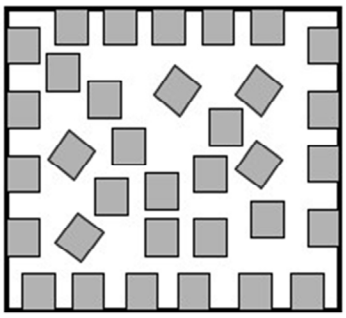
- Block #54 was selected as the approximate location for the first interview.
- The interviewer arrives. Now what?
- Number the houses and choose one randomly. Or you could choose one of the four sides randomly and number the houses on that side only, then choose one of them randomly.



Selecting the First House: Crowded Blocks or Districts

- Block #9 was selected as the approximate location for the first interview.
- The interviewer arrives. Now what?
- Divide the block into four quadrants or sections. Choose one sector randomly. Count the number of houses in that sector and select one house randomly.

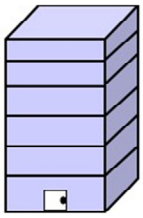
Block #9



Selecting the First House: An Apartment Building

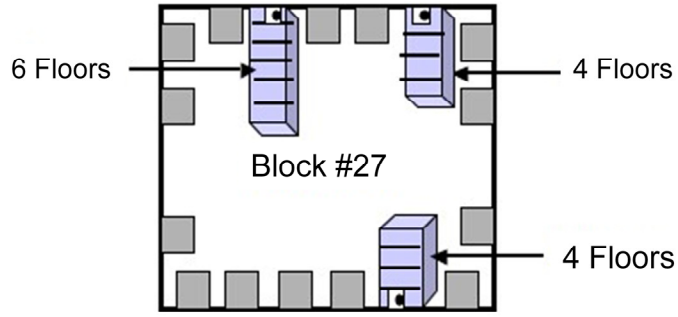
- Building #73 was selected as the approximate location for the first interview.
- It is an apartment building.
- The interviewer arrives. Now what?
- Number the floors of the apartment building.
  - Choose one floor randomly.
  - Count the number of doors on that floor. Choose one door randomly.
  - If you cannot use the rule of "go to the closest door" since all doors are next to each other or are equally close, then choose one rule before beginning your search, such as "go to the right."
  - Similarly, if you find no one on the floor you can interview, choose one rule before beginning your search, instructing you to go up one floor or down one floor.

Building #73



Selecting the First House: Apartments and Houses Together

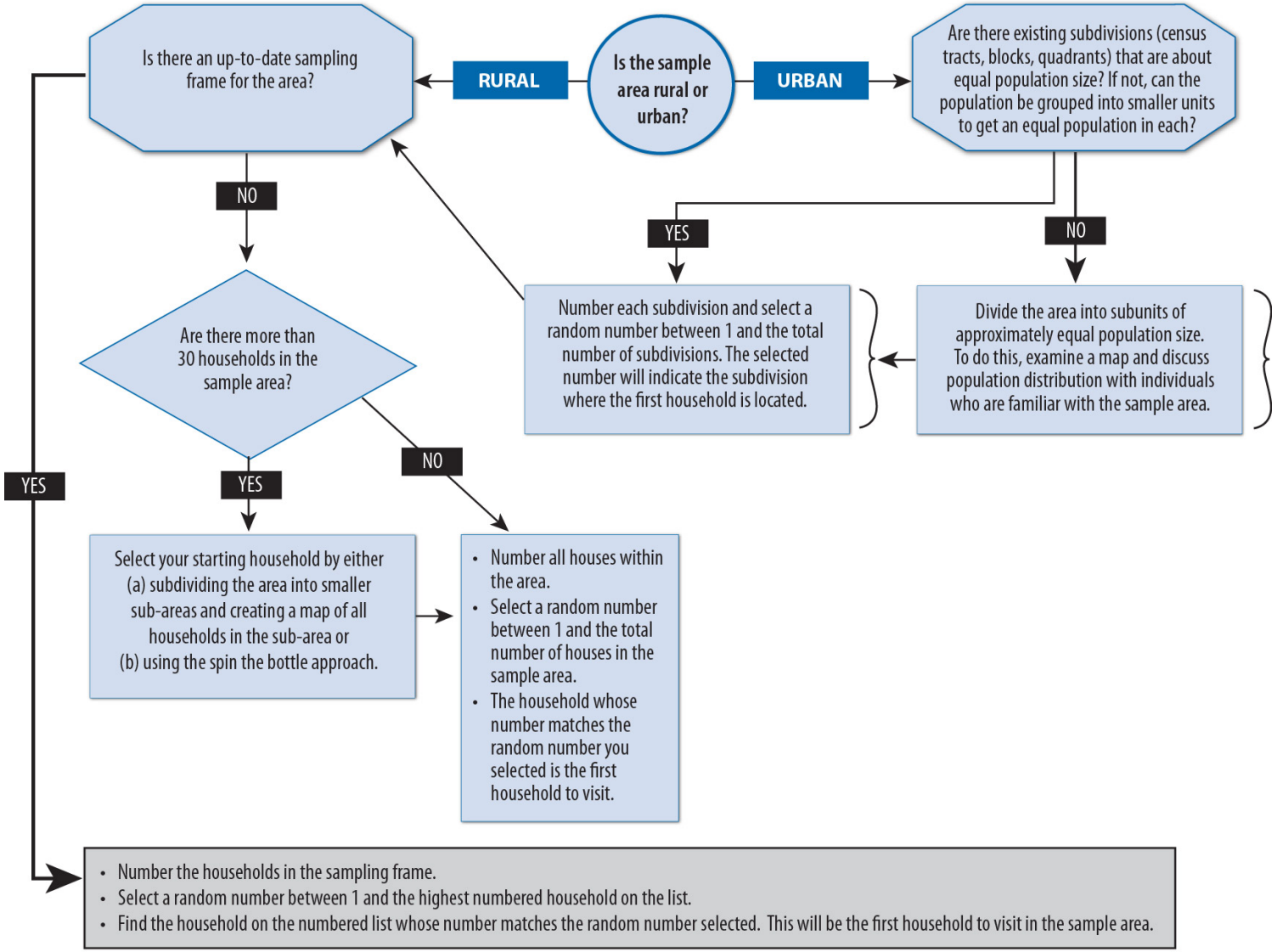
- Block #27 was selected as the approximate location for the first interview.
- There are both apartments and houses in this block.
- The interviewer arrives. Now what?
- In complex situations like this one, the simplest solution may be to count all of the houses/doors in the selected block and choose one randomly. You could estimate the number of doors in an apartment building by counting the number of doors on one floor and then multiplying this number by the number of floors in the building.



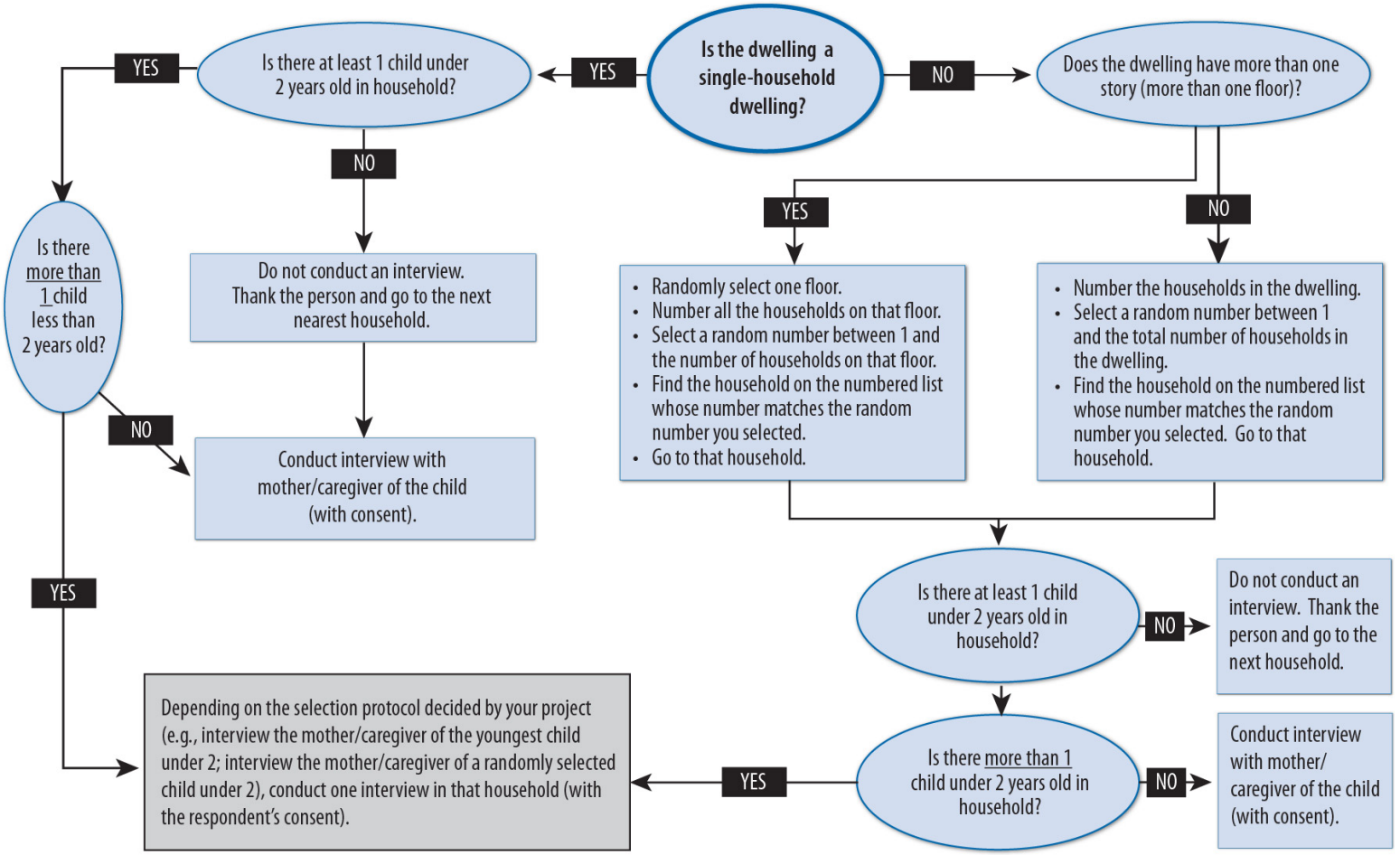
Steps for Numbering Households in Interview Locations with More Than 30 Households

1. Subdivide the community into two or more equal sections (you do not need to create a map of the entire interview location but should be sure that subdivisions are roughly equal in size).
2. Select one of these sections at random.
3. If selected area is still too large, subdivide it again into two or more equal sections, number each section, and select one section at random.
4. Continue until you have one small section with 15–25 households.
5. Draw a map of the section with the help of an informant.
6. Number the households in this section on this map (you only need to count the houses in the selected section).

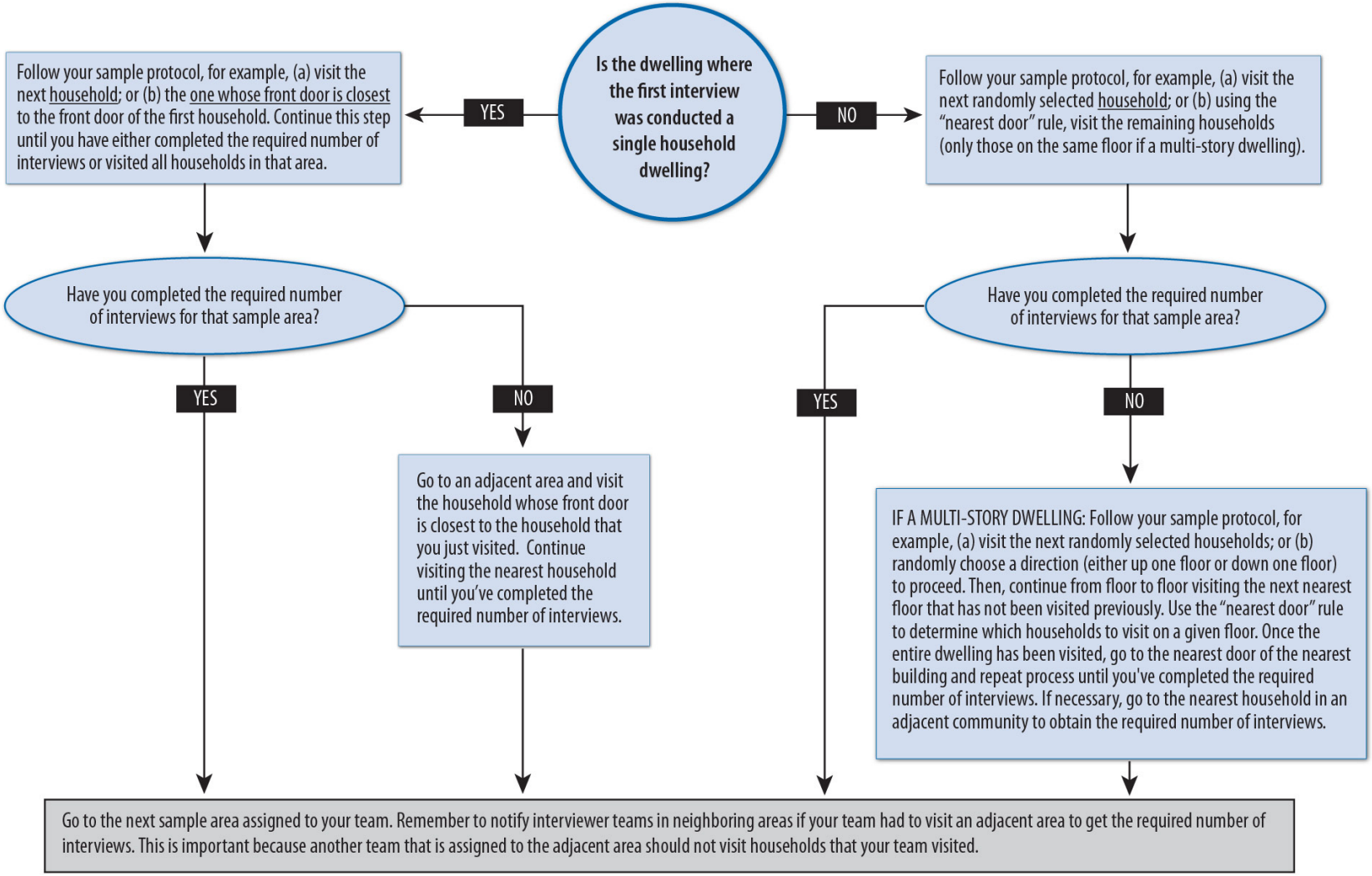
Selecting the First Household in an Interview Location

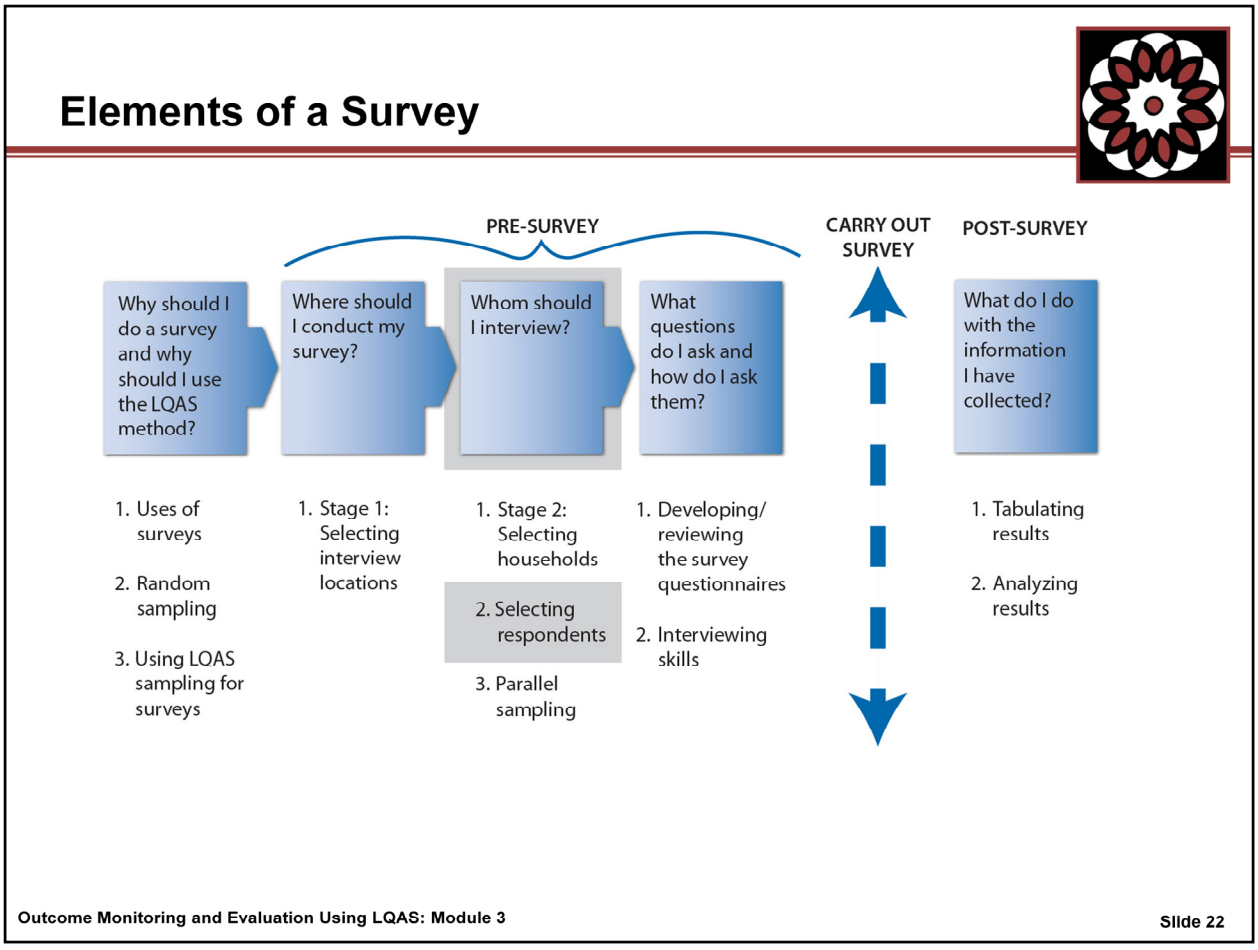


Choosing Respondents Based on the Types of Dwelling



Selecting Subsequent Households in an Interview Location





The selection of a household does not guarantee that we will actually be able to conduct an interview there. It will depend on whether the household contains the type of person we want to interview. Below we look at the simplest case of having a single kind of respondent for one indicator. Later we will discuss the situation that is more common, in which we have more than one indicator and more than one kind of respondent. Here, we will continue with the example we have been using throughout this training: young people age 15-24. The questions you will ask concern their knowledge of how to prevent the transmission of HIV.

The question we address now is how to select them and what to do if there is no one like this in the household.

Rules for 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* of the household where you currently are, and check at this "next nearest" household. *Continue* this process *until* you find the respondent type you are looking for. If two households are equally near, then choose the one with the closest door. Otherwise, choose one randomly.
Lives at that household **BUT** is absent and far away (*more than* 60 minutes 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. If you do find this person, *interview him or her* if he or she consents.

* Household = Group of persons who share the same kitchen or hearth; a group of persons who eat from the same cooking pot.

Hopefully these rules are clear. One thing that is not stated here is what to do if more than one person in the household has the characteristics of the respondent you are looking for. In that case you should use a random process to identify which person to interview. You can flip a coin if there are two or use a random number table to select one respondent. Once again, the process should be random. This means that even if the person who meets you at the door is the kind of respondent you are looking for, you should find out if there are others who fill the criteria and randomly select between them.

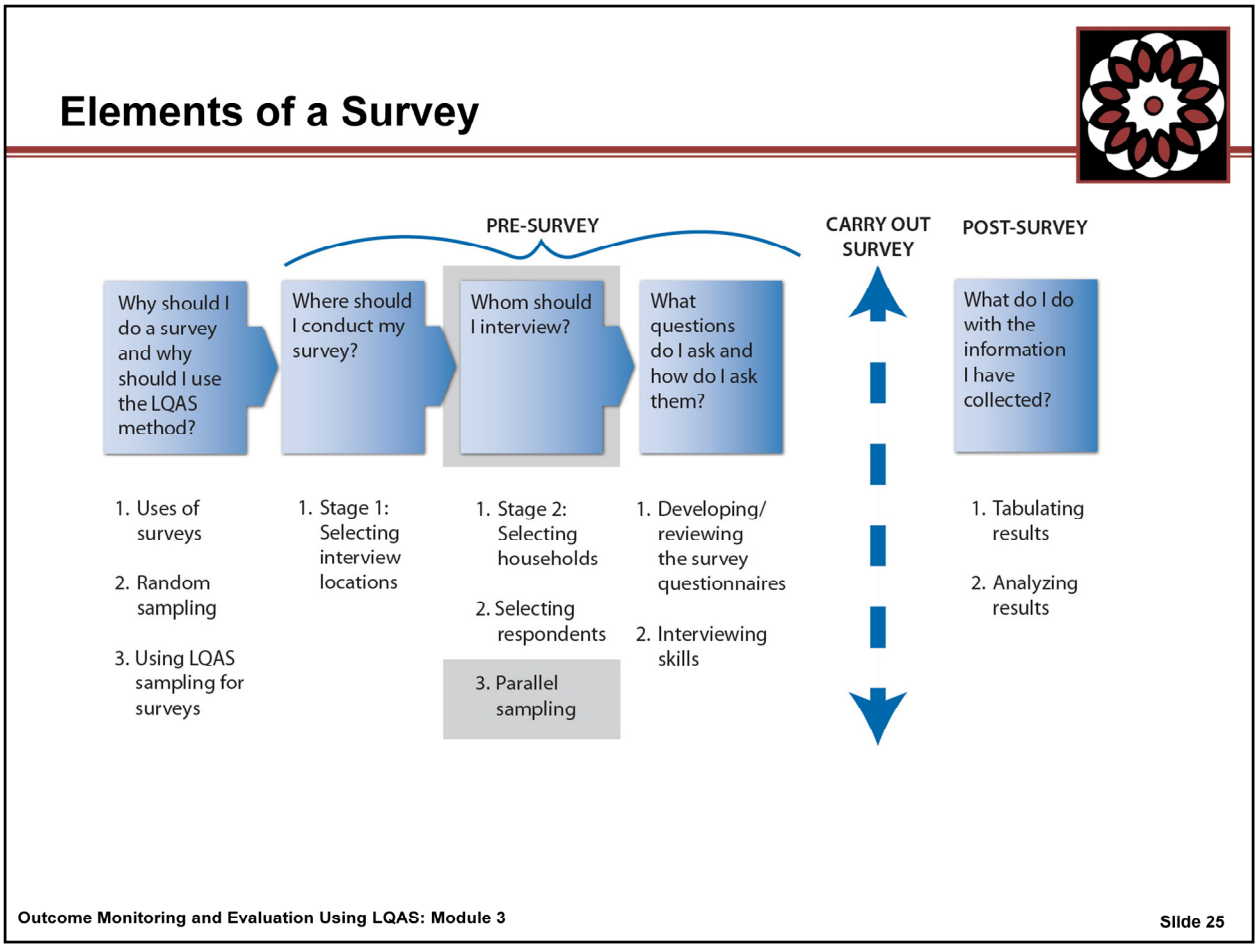
The point of using 60 minutes is to encourage interviewers not to give up too easily. Failing to interview people from the household who are not there at that minute—but available at a close distance—can introduce biases that are hard to quantify. The point here is to work hard to find the individual.

These are the same rules you would use if you were doing a 30- cluster survey. The only difference would be that once you interview a respondent, instead of moving on to the next interview location, you would go to the next household to interview the next respondent in the cluster. You would continue this until you have completed all the interviews for that cluster. At each household in the cluster, the same rules as these would apply. After each completed interview, you can go to the next nearest household from the front entrance of the household where you just completed the interview.

Activity: Selecting Respondents

- Each participant has been given a numbered scenario card representing a unique household that may or may not have an eligible respondent.
- Each participant will hold the scenario card and position themselves as households in the community.
- One participant will volunteer as an interviewer and will use a table of random numbers to select the first household.
- The person holding the scenario card for the selected household will read out loud the scenario described. The 'interviewer' will decide whether there is an eligible respondent.
  - If not, the interviewer moves to the next household whose entrance is nearest (and continues doing so till s/he finds an eligible respondent).
  - If so, the person who was holding this scenario card becomes the new interviewer and randomly select the next household.

Parallel Sampling



Throughout all of our discussions of LQAS to this point, we have mostly used a single indicator to illustrate how the approach works. We have done this to keep things simple. However, as we know, programs rarely concern a single issue and rarely have one group of interest in terms of the outcomes sought. Here, we have been using the example of the percentage of young people age 15-24 who know three ways to prevent HIV transmission.

While this indicator is typical of an HIV prevention program, there are likely to be other groups we want to work with as well. For example, we may want to ensure that pregnant women receive counseling and testing for HIV or that people living with HIV receive a minimum package of services or that orphans and other vulnerable children receive care.

As a result, when we monitor and evaluate our efforts at the population level, we will probably want to survey two groups (or even more, depending on the program).As a result, when gathering information from respondents about knowledge, practices, or service coverage, 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 describe the procedures for this in detail in this section.

Parallel Sampling (Indicator-Specific Sample Groups)

- In LQAS each indicator of interest to the project must have a complete sample (of 19) in each supervision area if we are to be able to pose the pass/fail judgment for the supervision area on the indicator
- At each interview location, it is unlikely that just one respondent will be eligible to answer questions relating to all the indicators of interest to a project.
- Since eligible respondents may differ from one indicator to another, relevant questions are grouped into questionnaires for the appropriate eligible respondent groups.
- At each sampled interview location, each type of questionnaire will be administered to only one eligible respondent.
- This approach is called “parallel sampling,” since each questionnaire with its eligible respondent group (“sub-sample”) actually constitutes a complete survey on a set of indicators for the project.

Let’s suppose that in addition to wanting to prevent the transmission of HIV among young people, we also want to ensure that all pregnant women are counseled concerning HIV and tested for HIV during prenatal visits. We have looked at the first case. For the second case, we will want to interview women who gave birth in the past year (who have a child less than a full year old) so we can ask them what happened throughout their pregnancy in terms of counseling and testing received.

Right away, we can see that if we select a household (as we did in the previous session), we actually may need to talk to two different people. At the same time, it is possible that the same person will be young (age 15-24) and recently pregnant. In the first case, we would not administer both surveys to one person, but in the second case, we could. We will see how this works below. The main point is that we must leave our interview location with 19 completed surveys for each survey we produce, based on the indicators of interest to us.

Examples of Indicators and Eligible Sub-Samples

Group of Indicators of Interest to the Project in the Survey	Eligible Respondent Groups
Exclusive breastfeeding rates	Mothers of children 0–5 months
Pre/postpartum care, including birth attendance, essential newborn care, pre/postpartum visits, vitamin A, TT, and iron folate	Mothers of children 0–11 months
Immunization and vitamin A, infant and young child feeding	Mothers of children 12–23 months
Malaria, illness management (diarrhea, fever, ARI), anthropometry	Mothers of children 0–59 months
Family planning, including CPR, needs satisfied, birth spacing; as well as reproductive health	Women of reproductive age (15–49 years)

These examples are not for HIV/AIDS programs but provide examples from other typical health programs that illustrate why parallel sampling may, at times, be necessary. They are provided as examples only.

Steps for Carrying Out a Parallel Sample

1. Determine the number of sample groups based on your program goals and related indicators.
2. Develop (use) a single questionnaire for each sample group.
3. Select interview location (already practiced).
4. Select the first household in a location (already practiced).
5. Select an appropriate respondent for at least one questionnaire in a household.
6. Select an appropriate respondent for the second (and subsequent) questionnaire in the same household,

   OR

   Select the next household looking for an appropriate respondent for the second (or subsequent) questionnaire.

Note that the application in one household of several questionnaires, each one collecting different information, **does not introduce bias into the sample**. This is because the household was selected randomly as required by LQAS and because a respondent was only asked once about questions for each indicator.

The main point of parallel sampling is that you need to make sure that you have assessed whether at least one survey can be conducted in the household you selected. There are several possible scenarios:

1.You could determine that there is no one in the household who can respond to either survey. In our example, that means that there are both no young people age 15-24 and no women who have a child less than 1 year old.

2.You could determine that there is one person (or more) in the household who can respond to one survey but no one who can respond to the second one. If there are two or more people who can respond to one survey, you must randomly select one of them. For the second survey, you will need to go to the next closest household.

3.You could determine that there is one person (or more) in the household who can respond to one survey and another person (or more) who can respond to the second survey. In this case, your work in this interview location will be done.

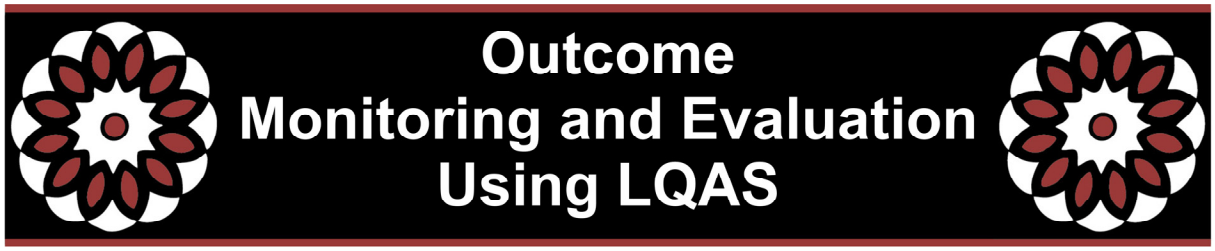
4.You could determine that one person is able to respond to both surveys. If that person is the only one who can respond to both, then you can administer both surveys to him or her. If there are others for one or both of the surveys, then in each case you will need to randomly select one person. This could mean that one person responds to both surveys. It could also mean that an individual, though eligible to respond to both surveys, ends up not responding to either.

Activity: Selecting Respondents in Parallel Sampling

1. Each participant has been given a numbered scenario card representing a unique household that may or may not have one or more eligible respondent(s).
2. Each participant will hold the scenario card and position themselves as households in the community.
3. One participant will volunteer as an interviewer and will use a table of random numbers to select the first household.
4. The person holding the scenario card for the selected household will read out loud the scenario described. The 'interviewer' will decide whether there is any or all eligible respondent(s).
    1. If none or not all, the interviewer moves to the next household whose entrance is nearest (and continues doing so till s/he finds all eligible respondent).
    2. If all, the person who was holding this scenario card becomes the new interviewer and randomly select the next household.

MODULE 4:

OBTAINING, TABULATING, AND USING SURVEY RESULTS



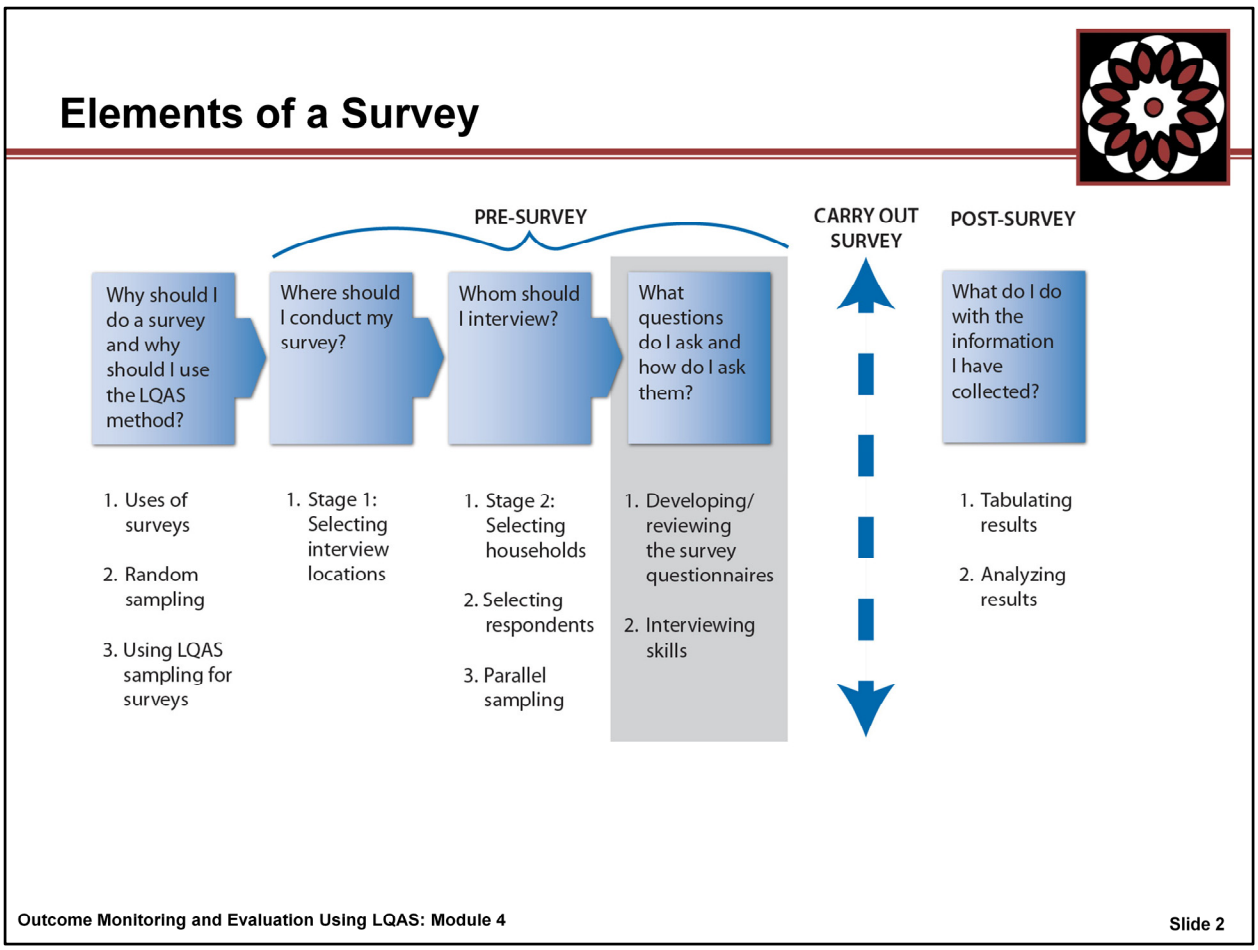
Module 4:

Obtaining, Tabulating, and Using Survey Results

Module 4: Objectives

1. **Practice** and **critique** each other's interview skills using a sample questionnaire
2. **Hand-tabulate** results for a single indicator from a "mock" survey
3. **Examine** the use of LQAS for baseline surveys and for routine M&E of progress towards targets
4. **Calculate** an indictor estimate for an entire program area
5. **Examine** the concept of precision and how confidence intervals portray it
6. **Evaluate** examples of how to calculate weighted estimates for that indicator in the same program area with confidence intervals

Questionnaires and Interviewing



Preparing to conduct a survey includes both preparing the actual questions to be asked and making sure that those asking them follow appropriate protocols that will assure accurate responses to questions.

A great deal of work has been done to develop questions and questionnaires for the indicators used in HIV/AIDS programming. We STRONGLY encourage you to use and adapt existing questionnaires rather than develop your own. Developing questions is a difficult task and not necessary for most indicators.

Activity: Review of Questionnaires

- Each participant will get the two short questionnaires that we will review together
- The different sections of the questionnaires:
  - Introduction and informed consent
  - Identification of eligible respondents
  - Respondent's characteristics (age, sex, marital status …)
- Different types of questions,
  - Open-ended
  - Closed (single response/multiple response)
  - Pre-coding
- Probes and prompts
- Skip patterns

Interview Etiquette

- Dress appropriately.
- Present official document/certificate from organization or project, if necessary.
- Do not enter the house unless you are invited.
- If you remain outside, do not ask for a chair; sit on the porch, steps, etc.
- Tell the interviewees how long the questionnaire will take.
- Do not accept lunch (unless it would be rude to refuse).
- Do not give gifts to interviewees.
- Thank interviewees at the end.

Others?

In structured surveys, it is critical to ask every respondent the same questions for a particular survey in the same way every time. If we are to be able to combine responses to draw conclusions about supervision area or an entire program area, it is critical that we maintain strict standards in terms of how we ask questions and probe responses.

In structured surveys, we do not want the interviewer to "get in the way" and influence the respondent's willingness to respond—and respond in an unbiased way—to questions. Thus, we must be careful about how we act and ask questions. First and foremost, we must know the survey instrument—the questionnaire—very well so that we are absolutely comfortable asking all the questions and marking the responses.

Effective Interview Techniques

1. Introduce yourself, your organization, and the purpose of the survey.
2. Maintain confidentiality:
   a. Do not interview the respondent in the presence of others.
   b. Explain that all answers will be kept confidential.
3. Ask questions exactly as written or with minor changes that were agreed upon during the training.
4. Wait for a response; be silent.
5. If the respondent doesn't understand or the answer is unclear (or he or she simply does not respond after a time), ask the question again.
6. Do not suggest—by tone of voice, facial expression, or body language—the answer you want.
7. Try not to react to answers in such a way as to show that you approve or disapprove.
8. If one answer is inconsistent with another, try to clear up the confusion. (This happens if two questions address similar information in two ways and is rare.)
9. Try to maintain a conversational tone of voice; don't make the interview seem like an interrogation.
10. For questions that allow respondents to list several responses, use neutral probes (e.g., anything more?) to see if they want to add anything.

By using these techniques, we can help ensure that all interviewees are treated the same and all questions are asked in the same way each time.

Activity: Interviewing Role Play

- We will demonstrate interviewing with a volunteer as the rest of the participants observe.
- The participants will form groups of three and, assuming roles of respondent, interviewer and observer, practice interviewing in turns. The observer shall give a debrief after each role play. This will continue until each participant has had a chance to play each of the three roles.
- We will then discuss in plenary what went well and areas needing improvement.

Field Practice

Activity: Preparation for Field Work

- We will practice what we have learned so far in a nearby community. The key skills we will practice are:
  - Random sampling to select households for LQAS
  - Respondent selection with parallel sampling
  - Interviewing using the sample questionnaires
- The participants will form teams of three (or six).
- Each team will be taken to a community with a sampled interview location.
- The team will be expected to select the first household for interview, identify eligible respondents, and conduct the interview.
- Each team is expected to complete at least two questionnaire sets.
- Each team will record their experience on the above skill areas for sharing with the rest on return from the field.
- We will then discuss in plenary what went well and areas needing improvement.

Field Practice Debrief

Activity: Plenary Fieldwork Debrief

1. Participants to sit in their field teams
2. Get the participants to mix and form new discussion groups such that each discussion group has a representative from each field team
3. Each discussion group to discuss their experience and challenges
4. Each discussion group to share with the plenary experiences and challenges that were common in the field teams

Randomness Reminder

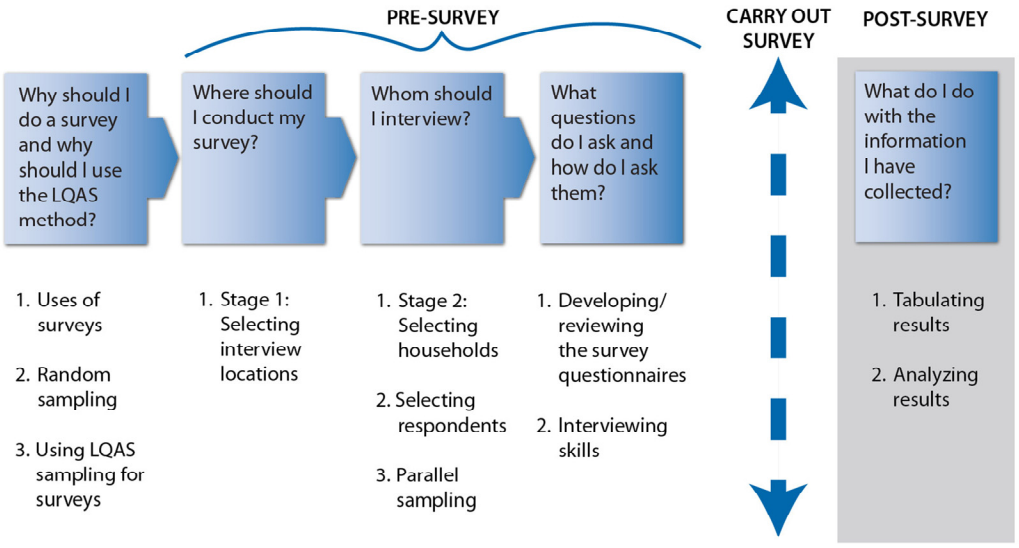
A commitment to random sampling means that you will—

- Randomly select communities in which surveys are to be conducted
- Randomly select subdivisions within a community after carefully mapping the community (if necessary)
- Randomly select a starting household after mapping the randomly selected subdivision
- Randomly select a respondent in the randomly selected household if there is more than one eligible respondent there
- Respect the 60-minute rule and going to the next nearest household
- Randomly select all respondents in a parallel sampling situation.

This slide summarizes all the points at which random selection can break down. We must make sure that random processes are maintained at each step.

Tabulation Results from a Single Supervision Area

Elements of a Survey



We are going to do tabulation by hand so we can get a feel for all the elements in this first step of analyzing data. Tabulation is a recording of information necessary to assess the indicators of interest. One can do it by computer but we will do it by hand to make sure all the elements are clear.

Result Tabulation HIV Transmission for a Supervision Area

Result Tabulation HIV Transmission for a Supervision Area: Young People Age 15–24

Supervision Area: _____ **Supervisor:** _____ **Tabulation Date:** _____

Correct = 1 **Incorrect = 0** **Missing = X**

#	Question/Indicator	Correct Response Key	Interview/Questionnaire Number																			Total Correct in SA	Total Sample Size (all 0's and 1's)
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
2	Have you ever heard of an illness called AIDS?	1. Yes = 1 2. No = 0—If no, then questions 3–5 and the Indicator = 0 also																					
3	Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?	1. Yes = 1 2. No = 0 8. DK = 0																					
4	Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?	1. Yes = 1 2. No = 0 8. DK = 0																					
5	Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all?	1. Yes = 1 2. No = 0 8. DK = 0																					
I	Knows three ways to prevent the transmission of HIV.	If all questions 2 through 5 are 1 (Yes) = 1 If any of questions 2 through 5 = X (Missing) and all others are 1 (Yes) – X Otherwise = 0																					

Outcome Monitoring and Evaluation Using LQAS: Module 4 **Slide 11**

Notice the following about this sheet:

1. This sheet is used to capture information directly from the surveys for a single supervision area. You would, therefore, have a single sheet like this one for each supervision area. You should also have a separate sheet or section of a sheet for each indicator. Since we only have one indicator of interest, ours is simpler. The second sheet for the second indicator is given below. If there were to be more than one indicator on this sheet, you would change the title to reflect the response group (young people age 15-24 in this case).

2. Notice that the second line allows you to provide information on the supervision area and the date of the tabulation.

3. The third line reminds you what numbers to put in each column (interview or questionnaire), and here you have three choices. If you put a number 1 for a question on a given questionnaire, it means the respondent answers “correctly” meaning they give you an answer that indicates they know the information, practice the behavior you want them to practice, or are covered by the service you want them to be covered by. If you put a 0 (zero), it means they do not know, do not practice, or are not receiving that service. You would place an X only if there was simply no answer given to a question. This should be a very rare thing, and it usually only happens because an interviewer forgets to circle an answer. This is different from skipping an answer. You could also have a code for skipped responses, but we are going to code them as 0 in this case, as you will see.

Result Tabulation HIV Transmission for a Supervision Area (continued)

Result Tabulation HIV Transmission for a Supervision Area: Young People Age 15–24

Supervision Area: _____ **Supervisor:** _____ **Tabulation Date:** _____

Correct = 1 **Incorrect = 0** **Missing = X**

#	Question/Indicator	Correct Response Key	Interview/Questionnaire Number																			Total Correct in SA	Total Sample Size (all 0's and 1's)
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
2	Have you ever heard of an illness called AIDS?	1. Yes = 1 2. No = 0—If no, then questions 3–5 and the Indicator = 0 also																					
3	Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?	1. Yes = 1 2. No = 0 8. DK = 0																					
4	Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?	1. Yes = 1 2. No = 0 8. DK = 0																					
5	Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all?	1. Yes = 1 2. No = 0 8. DK = 0																					
I	Knows three ways to prevent the transmission of HIV.	If all questions 2 through 5 are 1 (Yes) = 1 If any of questions 2 through 5 = X (Missing) and all others are 1 (Yes) – X Otherwise = 0																					

Continued…

4. Now let's go column by column. The first column indicates the question number corresponding to the tabulation and refers you to the questionnaire. An "I" means that it is an indicator constructed from the results of one or several questions. Though we tabulate individual questions (and may analyze their results), we are most interested in this indicator. Notice in our example that the first question on age is not included. This is because, if we have done our work correctly, everyone who responded to our questions was of an appropriate age, and we do not need to analyze it further. However, you probably will want to analyze the age distribution of respondents, and computer entry of all surveys can help with that.

5. The next column actually gives the question (or if it is long, a part of the question). In the case of the indicator, it defines the indicator of interest.

6. The next column is very important and tells you the information the tabulation team needs to enter: either 1, 0, or X. Since these are written in a shorthand form, we should look at them carefully. Let us see some examples. For Question 2, the "correct" response key means that if the person said "yes" and 1 was circled on the survey, you would write a 1 in the appropriate column. This is the "correct" answer in that we wanted them to have heard of HIV. If, on the other hand, the respondent said "no" and 2 was circled, you will enter a 0 in the appropriate column. Notice also that if they answered "no" to Question 2, that you would have skipped all the other questions. In that case, you will automatically place 0 in all the remaining rows--including the indicator row for this questionnaire.

Result Tabulation HIV Transmission for a Supervision Area (continued)

Result Tabulation HIV Transmission for a Supervision Area: Young People Age 15–24

Supervision Area: _____ **Supervisor:** _____ **Tabulation Date:** _____

Correct = 1 **Incorrect = 0** **Missing = X**

#	Question/Indicator	Correct Response Key	Interview/Questionnaire Number																			Total Correct in SA	Total Sample Size (all 0's and 1's)
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
2	Have you ever heard of an illness called AIDS?	1. Yes = 1 2. No = 0—If no, then questions 3–5 and the Indicator = 0 also																					
3	Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?	1. Yes = 1 2. No = 0 8. DK = 0																					
4	Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?	1. Yes = 1 2. No = 0 8. DK = 0																					
5	Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all?	1. Yes = 1 2. No = 0 8. DK = 0																					
I	Knows three ways to prevent the transmission of HIV.	If <u>all</u> questions 2 through 5 are 1 (Yes) = 1 If any of questions 2 through 5 = X (Missing) and all others are 1 (Yes) – X Otherwise = 0																					

Outcome Monitoring and Evaluation Using LQAS: Module 4 **Slide 11**

Continued…

We could use an S for skip, but we want to simplify our work. We are saying that if they have not heard of HIV, then they do not know how to prevent it.

Notice Question 3. Here, if the person responded “yes” and 1 is circled on the questionnaire, we place a 1 in the appropriate column. If they said “no” or said they “don’t know” we consider that to be wrong in both cases and place a 0 in the appropriate column. Questions 4 and 5 are coded the same way.

Finally, notice the indicator. This one is a little tricky, so let’ go through it. It says this: If Questions 2-5 (above) are all 1, then you should place a 1 in this row also. Next, if all the rows for Questions 2-5 (above) contain an X because the interviewer forgot to fill in a question properly, then we have to consider that the indicator also is “missing.” This means we cannot analyze it, and our sample goes from 19 to a lower number. Finally, in every other case, we should place 0 for the indicator. Thus, if there are any 0’s above in any row--even if there is also an X in another place--we should place 0 here. Thus, if any of the questions have 0 next to them, we judge that the person does not know three ways to prevent the transmission of HIV.

Result Tabulation HIV Transmission for a Supervision Area (continued)

Result Tabulation HIV Transmission for a Supervision Area: Young People Age 15–24

Supervision Area: _____ **Supervisor:** _____ **Tabulation Date:** _____

Correct = 1 **Incorrect = 0** **Missing = X**

#	Question/Indicator	Correct Response Key	Interview/Questionnaire Number																			Total Correct in SA	Total Sample Size (all 0's and 1's)
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
2	Have you ever heard of an illness called AIDS?	1. Yes = 1 2. No = 0—If no, then questions 3–5 and the Indicator = 0 also																					
3	Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?	1. Yes = 1 2. No = 0 8. DK = 0																					
4	Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?	1. Yes = 1 2. No = 0 8. DK = 0																					
5	Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all?	1. Yes = 1 2. No = 0 8. DK = 0																					
I	Knows three ways to prevent the transmission of HIV.	If all questions 2 through 5 are 1 (Yes) = 1 If any of questions 2 through 5 = X (Missing) and all others are 1 (Yes) = X Otherwise = 0																					

Continued…

The issue of "missing" can be confusing, so let's look at it quickly. What we are saying here is we can judge whether a person does not know three ways to prevent HIV even if some data are missing if at least one other question has a 0. However, we can never know if a person does know three ways if there is one missing response because that missing data could have been 0--we cannot know.

7. The next 19 columns are where you actually enter 1, 0, or X for each question and for each survey numbered 1-19. This reminds us that, before data entry can begin, we must clearly number each survey consecutively.

8. The column after 1-19 is where you add up the 1's from the previous 19 columns. This represents the total number of "correct" responses (this is where we start getting excited because we can actually see the results).

9. F inally, the last column provides us with our true sample size for a given question or indicator. Here, we total up the 0's and 1's from the 19 columns. Normally, this should be 19, but if there are any X's, then the number would be fewer as we have seen.

Result Tabulation Counseling and Testing for a Supervision Area

Result Tabulation Counseling and Testing for a Supervision Area: Mothers of Children Less than 1 Year Old

Supervision Area: _____ **Supervisor:** _____ **Tabulation Date:** _____

Correct = 1 **Incorrect = 0** **Missing = X**

#	Question / Indicator	Correct Response Key	Interview/Questionnaire Number																			Total Correct in SA	Total Sample Size (all 0's and 1's)
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
2	Did you see anyone for antenatal care before the birth of (NAME OF CHILD)?	1. Yes = 1 2. No = 0—If no, then questions 3–6 and the Indicator = 0 also																					
3a	During any of the antenatal visits for that birth, did anyone talk to you about: ▪ Babies getting the AIDS virus from their mother?	1. Yes = 1 2. No = 0 8. DK = 0																					
3b	During any of the antenatal visits for that birth, did anyone talk to you about: ▪ Things that you can do to prevent getting the AIDS virus?	1. Yes = 1 2. No = 0 8. DK = 0																					
3c	During any of the antenatal visits for that birth, did anyone talk to you about: ▪ Getting tested for the AIDS virus?	1. Yes = 1 2. No = 0 8. DK = 0																					
4	Were you offered a test for the AIDS virus as part of your antenatal care?	1. Yes = 1 2. No = 0																					
5	Were you tested for the AIDS virus as part of your antenatal care?	1. Yes = 1 2. No = 0																					
6	Did you get the results of the test?	1. Yes = 1 2. No = 0																					
I	Women counseled and tested.	If all questions 2 through 5 are 1 (Yes) = 1 If any of questions 2 through 5 = X (Missing) and all others are 1 (Yes) = X Otherwise = 0																					

This indicator is a coverage indicator whereas the other one was a knowledge indicator. Notice how Question 3 has been split into three. Note that, for the indicator to be 1, the response to all questions must be “yes”.

The logic of this indicator tabulation sheet is exactly like the first one.

LQAS Table Decision Rules

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 can not 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.

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

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

Let's assume that this survey is being conducted after 2 years of program implementation, and the program set a target for each SA that each one should be at 75% coverage at this time.

What is your decision rule?

Summary Tabulation HIV Transmission Knowledge for a Program

Summary Tabulation HIV Transmission Knowledge for a Program: Young People Age 15–24

Program Name: _____ **Summary Tabulation Date:** _____

#	Question/Indicator	Total Correct in Each SA (with decision rules for target coverage [left] and average coverage [right])					Total Correct in Program	Sample Size in Each SA					Total Sample Size in Program	Average Coverage= Total Correct/ Sample Size	Target Coverage
		A	B	C	D	E		A	B	C	D	E			
2	Have you ever heard of an illness called AIDS?														
3	Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?														
4	Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?														
5	Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all?														
I	Knows three ways to prevent the transmission of HIV.														

After the tabulations for each supervision are complete you can transfer the information from the individual supervision area tabulation sheets to this summary tabulation form.

The first row tells us which indicator this refers to, and even though we are going to tabulate summary results for individual questions and a single indicator, we are only going to assess program-level results for the indicator.

The second row is summary information. Now let's look at it column by column. The first two columns are identical to what we saw on the individual supervision area tabulation sheets--the question number and question or indicator. The next set of columns has several features.

First, the top box for each row is where you transfer the information from the "Total Correct in SA" column from the individual supervision area tabulation forms. The two smaller boxes below the main box are where you can place your decision rule based on your target. So, on the left box we would place the decision rule for our target. Note that if there had been one or more supervision area that had missing data and had, say, only 17 or 18 respondents, we would need to find the decision rule for those sample sizes rather than for 19. The results would be less precise, but that is what we would have to do.

Summary Tabulation HIV Transmission Knowledge for a Program (continued)

Summary Tabulation HIV Transmission Knowledge for a Program: Young People Age 15–24

Program Name: _____ **Summary Tabulation Date:** _____

#	Question/Indicator	Total Correct in Each SA (with decision rules for target coverage [left] and average coverage [right])					Total Correct in Program	Sample Size in Each SA					Total Sample Size in Program	Average Coverage= Total Correct/ Sample Size	Target Coverage
		A	B	C	D	E		A	B	C	D	E			
2	Have you ever heard of an illness called AIDS?														
3	Can people reduce their chance of getting the AIDS virus by having just one uninfected sex partner who has sexual intercourse with no other partners?														
4	Can people reduce their chance of getting the AIDS virus by using a condom every time they have sex?														
5	Can people reduce their chance of getting the AIDS virus by not having sexual intercourse at all?														
I	Knows three ways to prevent the transmission of HIV.														

Outcome Monitoring and Evaluation Using LQAS: Module 4 **Slide 14**

The next column, "Total Correct Program," is where you would sum up the total correct from each SA to get a total correct for the entire program.

The next set of columns is where you would enter the actual sample size (remember problem of "Missing" that can occur) for each supervision area.

The next to last column "Average Coverage=Total Correct/Sample Size" is something we will discuss in more detail in the next session, but it is our initial program-wide estimate of knowledge (in this case) for this indicator (or for a specific question). This is important information to have since many organizational leaders and donors are most interested in the program-wide knowledge, behavior, or coverage level.

The final column "Target Coverage" is where we can place the target we will be using to analyze results. As you can see, we don't set coverage targets for individual questions—though we could. For example, we could say that we want 90% of young people to have heard of HIV. In our example, we will not do that and merely set a coverage target for the final line—our key indicator.

Activity: Tabulation of a Single SA in Groups

- Participants will form five groups.
- Each group will be provided with:
  - A set of 19 completed questionnaires
  - A tabulation sheet
- The group will designate three people to:
  - Read the responses from the questionnaire
  - Record responses on the tabulation sheet
  - Verify the work of these two people
- Each group to complete the tabulation up to the “total correct” and “total denominator.”
- There will be plenary discussion of the results using hypothetical set target.
- Finally the plenary will be taken through summary tabulation for a program (summarizing results from all SAs of the program).

Different Uses of LQAS

Summary Tabulation HIV Transmission Knowledge for a Program: Young People Age 15–24

Indicator	Correct Responses by SA (decision rule target-left, decision rule average coverage-right)										Total Correct	Sample Size by SA					Total Sample	Average Correct	Target
	A		B		C		D		E			A	B	C	D	E			
Knows three ways to prevent the transmission of HIV	12		9		16		11		14		62	19	19	19	19	19	95	62/95= **65.3%**	75%
	12		12		12		12		12										

65.3% → Round up to 70%

Decision Rule = 11

Up until now, we have been using a simulation that is assuming that we have set a target based on some previous information, carried out the program, and are now coming back to assess our progress toward the target. In our example, the target is 75% of young people age 15-24 who will know three ways to prevent the transmission of HIV.

We can also use LQAS at baseline, and though our interpretations are a bit different, we can still draw two kinds of conclusions. Let's continue using the same example we have been using. Let's say that instead of finding these results at some point after program implementation, we found them at baseline before the program began. What would be different? Well, first, we would not yet have a target (like 75%) because we would not even know what the level was. However, we can use LQAS to assess whether certain supervision areas are substandard in relation to the overall coverage. Let's see how to do that.

Look again at the results from the survey, and let's assume it is now a baseline survey. We can use the "average coverage" estimate (the program-level result) to find a decision rule to assess whether any SAs are substandard in relation to the "average." In our example, we found that the average was 65.3%. We can round that number up to the nearest 5 percentage points and find a decision rule at that level. In our case, that is 70%. The reason we round up is that LQAS (used in the way we use it here) is designed to find substandard supervision areas. By rounding up, we require a higher number of "correct" responses, and we are more likely to correctly identify substandard supervision areas (vis-à-vis the average). If we round down, more truly substandard supervision areas might be wrongly classified as not being substandard.

Summary Tabulation HIV Transmission Knowledge for a Program: Young People Age 15–24

Indicator	Correct Responses by SA (decision rule target-left, decision rule average coverage-right)										Total Correct	Sample Size by SA					Total Sample	Average Correct	Target
	A		B		C		D		E			A	B	C	D	E			
Knows three ways to prevent the transmission of HIV	12		9		16		11		14		62	19	19	19	19	19	95	62/95= **65.3%**	75%
	12	***11***	12	***11***	12	***11***	12	***11***	12	***11***									

65.3% → Round up to 70%

Decision Rule = 11

Now, we can see what to place in the small box below the LQAS result to the right.

It is important to note that what we learn at baseline can and should influence our program implementation decisions. For example, if we find that 65% of young people know three ways to prevent HIV transmission (as we did in our example), we may decide that we do not have to give as much emphasis to educating young people as we would if we had found that only 40% had this knowledge. We are not discussing target setting in this training, but baseline surveys are designed to help set targets and determine program priorities.

Throughout this training, we have been examining only one, or at times two, indicators. Most HIV/AIDS programs are tracking more than just two indicators, and so you are likely to be making decisions about several indicators at once. We believe that understanding the steps we have gone through for one or two indicators will enable you to track progress and identify supervision areas in need of immediate attention in programs in which there are several indicators. Our hope, however, is that you will limit the number of indicators you are tracking to enable you to focus on the most important issues for your program.

Confidence Intervals and Weighting

Anytime we use a sample, the estimate we produce—65.3% in the example we have been using--will likely be close to but different from the true percentage in the population. Recall that we sample because it saves time and money, but the tradeoff is that we have some uncertainty about how well the result we get represents what is true about the population.

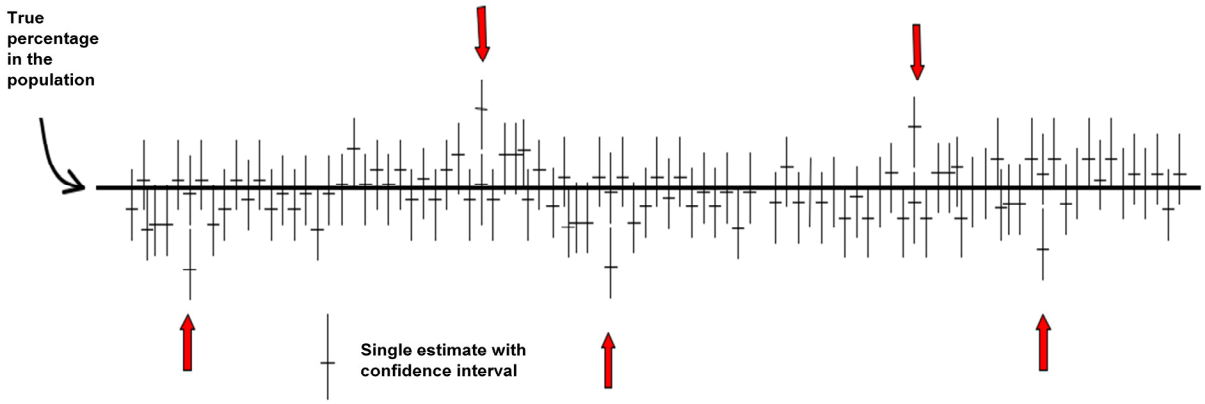
Unless we are willing to do a full census of a population, we will never know for sure the true percentage. However, if we use random sampling processes at every stage of our sample, we not only can produce an estimate but also express how certain we are about the result.

We do this by calculating something you may have heard of—a confidence interval. You may have heard someone talk about 95% confidence intervals. What does it mean?

Think about confidence intervals like this: They are intervals around the estimate (what we call the point estimate) coming from our sample for which we are 95% confident that the true population percentage is within the interval.

In other words, when we calculate and present the estimate, we are nearly certain (95% confident) that the true population percentage is within the interval.

Confidence Intervals



But what does it really mean to say we are 95% confident? We can think of it this way:

Let's pretend that instead of taking a single sample in our program area (5 samples of 19 is considered a single sample here), we were able to take 100 samples. Of course, we would never do that—but if we did, and used random sampling procedures, we could calculate a confidence interval each time. Each time, we would have a point estimate indicated in this picture by a horizontal line and an interval, indicated by the vertical line. The point estimate would be in the middle of the interval as it is here.

If we calculated 100 intervals, 95% of them would contain the true population percentage, and 5% of them would not—just as in the picture. Of course, we can only take one sample, and we can only calculate one interval, and we assume that our interval is 1 of the 95 that contain the interval. In that sense, we are 95% confident.

Here is the important point: Each time you provide a point estimate, you should also provide a confidence interval because that tells everyone how confident you are of the result. Of course, all of this assumes that you have sampled randomly at each stage as we have now learned how to do.

Calculating a 95% Confidence Interval

$$\text{Confidence interval} = p \pm 1.96 * \sqrt{((p * q)/n)}$$

1. **p** in this formula is the total correct from our survey (show the large format result) divided by the total sample. In other words, it is our coverage estimate or average coverage. We don't express it as a percentage but as a number—no 65.3% but rather as 0.653.
2. **q** is merely 1-p (or in our case, 0.347).
3. **n** is the size of our sample—in our case, n=95.
4. **1.96** is from a statistical table and we will not describe it further. You must only know to use this every time.

So what this formula says is—

1. Multiply p times q and divide that by n,
   - In our case, this is (0.653 * 0.347) / 95 = 0.0024.
2. Take the square root of this number (any calculator will do this).
   - In our case, this is SQRT 0.227 = 0.0489.
3. Multiply this result by 1.96.
   - In our case, this is 1.96 * 0.04885 = 0.0957.
4. To get an interval subtract this final result from p then add it to p.
   - In our case, 0.653 - 0.0957 = 0.557 or 55.7%.
   - In our case, 0.653 + 0.0957 = 0.748 or 74.8%.

Thus, our 95% confidence interval is 55.7 to 74.8%. We say we are 95% confident that the true population percentage of young people who know three ways to prevent HIV transmission is between 56 and 75% (approximately).

This interval may seem "wide" to you, and you can see from the formula that to make it "narrower" would require a larger "n"--a larger sample size. Unfortunately, increasing the sample size is costly and even raising it a great deal will not reduce the interval much. We will talk more about sample size in the next session, but you can see that a sample size of 95 (5*19) yields a confidence interval that is about plus or minus10 percentage points around the average coverage figure.

Weighting Examples

1	2	3	4	5	6	7	8	9	10
Supervision Area	**Sample Size** ss	**Correct Responses** c	**Mini Proportion** p=c/ss	**Population Size** n	**Weight** Wt=n/Total n	**Weight * Mini Proportion** Wt * p	q=1-p	p*q	**Weighted Standard Error** $Wt^2*((p*q)/n)$
A	19	12	0.632	6379	0.146	0.092	0.368	0.233	0.00026
B	19	9	0.474	10718	0.245	0.116	0.526	0.249	0.00079
C	19	16	0.842	7500	0.172	0.145	0.158	0.133	0.00021
D	19	11	0.579	9731	0.223	0.129	0.421	0.244	0.00064
E	19	14	0.737	9379	0.215	0.158	0.263	0.194	0.00047
	95	62	0.653	43707	1.000	0.640	0.347	0.227	0.00236

Mean **65.3%** **95% Confidence Interval** **55.7%** **74.8%**
Weighted Mean **64.0%** **Weighted 95% Confidence Interval** **54.5%** **73.5%**

Weighted Confidence Interval = p±1.96*(SQRT(Sum of Weighted Standard Error Column) 0.640±1.96(SQRT(0.00236))

1	2	3	4	5	6	7	8	9	10
Supervision Area	**Sample Size** ss	**Correct Responses** c	**Mini Proportion** p=c/ss	**Population Size** n	**Weight** Wt=n/Total n	**Weight * Mini Proportion** Wt * p	q=1-p	p*q	**Weighted Standard Error** $Wt^2*((p*q)/n)$
A	19	12	0.632	2379	0.054	0.034	0.368	0.233	0.00004
B	19	9	0.474	20718	0.474	0.225	0.526	0.249	0.00295
C	19	16	0.842	2000	0.046	0.039	0.158	0.133	0.00001
D	19	11	0.579	11731	0.268	0.155	0.421	0.244	0.00092
E	19	14	0.737	6879	0.157	0.116	0.263	0.194	0.00025
	95	62	0.653	43707	1.000	0.569	0.347	0.227	0.00418

Mean **65.3%** **95% Confidence Interval** **55.7%** **74.8%**
Weighted Mean **56.9%** **Weighted 95% Confidence Interval** **44.2%** **69.5%**

These tables provide two examples of weighting. The columns in each table calculate the same things and the only difference between the two is that in the second the population is distributed differently with very large and very small SAs (unlike the first in which they are similar). Note the red text in the tables.

Note that in both tables the mean and 95% confidence intervals are the same (65.3% and 55.7% to 74.8%). This is as one would expect since these confidence intervals are not weighted by the relative population in the SAs.

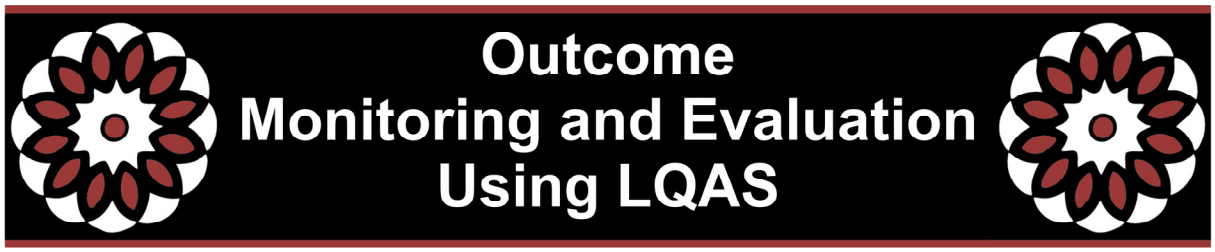
If you look at the weighted means you will notice some things. In the first table the weighted mean is a bit lower than the mean 64.0% to 65.3%. This may not be surprising because the SA with the lowest number knowing three ways to prevent HIV transmission—SA B—has the highest population which gives it more weight. Notice, however that the confidence intervals for both estimates largely overlap meaning that we have very little reason to think that weighting would change our conclusions about the overall coverage of the program.

In the second example the population disparities are much greater and here SA B and SA D—the SAs with the lowest numbers in the sample who know three ways to prevent HIV transmission—together have nearly 70% of the population. The weighted point estimate is very different from the unweighted one—56.9% to 65.3% but, again, the weighted confidence intervals still overlap quite a bit—with the weighted one quite a bit wider.

For details on the meaning of each column and how to fully calculate these see the *Rapid Health Surveys* reference in your resources.

MODULE 5:

SAMPLING DECISIONS, LOGISTICS, AND REPORTING



Module 5:

Sampling Decisions, Logistics, and Reporting

Module 5: Objectives

1. **Analyze** various sample-size options when conducting LQAS

2. **List** critical logistical issues that must be addressed in conducting population-based surveys

3. **Consider** elements of a report of a baseline and/or routine M&E results

Number of Supervision Areas and Sampling

Influence of Number of SAs on Program-Level Estimate Precision (assuming a sample size of 19 per SA)

Five Supervision Areas

Supervision Area	Sample Size	Correct Responses	
A	19	12	
B	19	9	
C	19	16	
D	19	11	
E	19	14	
Totals	95	62	
Mean		65.3%	
Confidence Interval		55.7%	74.8%
Minus/Plus		9.6%	

Six Supervision Areas

Supervision Area	Sample Size	Correct Responses	
A	19	12	
B	19	9	
C	19	16	
D	19	11	
E	19	14	
F	19	12	
Totals	114	74	
Mean		64.9%	
Confidence Interval		56.2%	73.7%
Minus/Plus		8.8%	

Eight Supervision Areas

Supervision Area	Sample Size	Correct Responses	
A	19	12	
B	19	9	
C	19	16	
D	19	11	
E	19	14	
F	19	13	
G	19	10	
H	19	14	
Totals	152	99	
Mean		65.1%	
Confidence Interval		57.6%	72.7%
Minus/Plus		7.6%	

Four Supervision Areas

Supervision Area	Sample Size	Correct Responses	
B	19	9	
C	19	16	
D	19	11	
E	19	14	
Totals	76	50	
Mean		65.8%	
Confidence Interval		55.1%	76.5%
Minus/Plus		10.7%	

Three Supervision Areas

Supervision Area	Sample Size	Correct Responses	
A	19	12	
B	19	9	
C	19	16	
Totals	57	37	
Mean		64.9%	
Confidence Interval		52.5%	77.3%
Minus/Plus		12.4%	

In the examples we have been providing, we have always had a program area with five supervision areas. In each of these five supervision areas, we proposed taking a sample of 19 for a total sample of 95. We see the value of these sample sizes but, of course, every program is different. Supervision areas should be based on either geographical or other considerations, such as language or ethnicity considerations, but the main point is that programs in each one are largely responsible for their own work and need to be able to make their own decisions. No matter how many supervision areas you have, a sample of 19 in each will allow you to identify substandard supervision areas that need immediate attention. This is the first and most important use of LQAS. Recall that increasing the sample size does not greatly improve ability to detect substandard areas unless you raise it a great deal.

However, we have also seen that programs usually want to combine the results of individual LQA samples from each supervision area in order to estimate the level of knowledge, practice, or coverage for the program as a whole. We talked about this and how to provide not only a point estimate but a confidence interval around each point estimate so that everyone will know the precision of the estimate.

This slide provides some examples of what happens to the precision of your estimates (the width of the confidence intervals) if you have fewer or more supervision areas in a program, and you take 19 from each. We constructed the examples so that the means are similar (around 65.5% or so). We can see that larger sample sizes increase precision of estimates (the confidence intervals are narrower). Even larger sample sizes do not increase precision a great deal. Smaller sample sizes (fewer supervision areas) lead to less-precise estimates.

Some Final Words on Sample Size…

There are other things to consider in determining sample sizes:

- If you wanted to compare a final evaluation with a baseline, you would need to determine a sample size based on how much change you expected to be able to carry out in pre- and post-intervention comparisons. We are not going to show you how to do that here.
- Also, if you wanted to marshal evidence that the changes in a program were "caused" by your program (in our approach, we assume that the program contributed to changes, but we cannot "prove" that), you would need a completely different pre- and post-intervention survey design with much larger sample sizes. Again, it is far beyond the purpose of this training to show how to do such studies.
- Sample size calculations can be complex, and we have merely tried to give you some rules of thumb to follow in determining what sample sizes might be useful to you and your program.

Activity: Review of Quality Checklists

- Participants will review their handouts and note issues that are unclear, or that are in their view very important.
- There will then be a plenary discussion during which participants can suggest additions to the quality checklists.

Quality Checklists

Pre-survey Checklist

Before the survey begins, be sure the following tasks have been completed:

1. Review the sampling frame (list of communities you selected from a list using probability proportionate to size) before designing the plan for data collection.
2. Count the questionnaires to be sure you have 19 for each respondent type (parallel sample!) and for each supervision area.
3. Number questionnaires 1 through 19 for each supervision area.
4. Review each one of the 19 questionnaires to make sure that they have the correct number of pages and they are securely stapled.
5. Verify questionnaires have no missing pages and are securely stapled together.
6. Review the materials checklist below. Be sure you have (or have decided you don't need) the following materials:

Materials Checklist

- ❑ 19 questionnaires for each respondent type + 2-3 extras of each for every SA
- ❑ Pencil
- ❑ Pencil sharpener
- ❑ Eraser
- ❑ Clipboard
- ❑ Day pack or bag to carry questionnaires and materials
- ❑ Random number tables
- ❑ Rules to select respondents in a household
- ❑ Raincoat
- ❑ Community maps or paper for making maps
- ❑ Any "questionnaire-specific" materials: condom brands, pictures of key informational brochures, etc.)

Data Collector's Checklist

After you are in the field, make sure you complete the survey in the following manner:

1. If a community census is available, number households and randomly select a starting household (and proceed as in step 6. below).
2. If no community census is available, update community maps, as needed, before selecting starting household(s), identifying all houses in the community. If no map is available, make one, being sure to include landmarks and showing the relative number of houses in each section of the community.
3. If the community is small, e.g., less than 30 houses, number all houses. Randomly select a starting household.
4. If the community is large, e.g., more than 30 houses, divide into sections (each section with a similar number of houses):
    - Number each section
    - Randomly select one of the community sections (if you have divided the community into 3 sections, select a random number between 1 and 3)
    - Go to the selected section to confirm the number of houses and the location of each house (and, if necessary, update the community map); if the section is large, subdivide it into subsections and randomly select one (and repeat this process until you get a subsection with 30 or fewer houses)
    - Number on the map each house in the section or subsection selected
    - Randomly select one house
5. If it is very difficult to divide the community or a section of it into sections, then—
    - Ask a respondent to take you to a place where exactly 50% of the houses are in front of you, 50% of the houses are behind you, 50% are to the right and 50% are to the left;
    - Number these 4 sections;
    - Choose one randomly;
    - Go to that section and repeat the procedure until you can see a manageable number of houses you can easily count; and
    - Select one of those houses randomly.
6. Go to the selected house to begin interviewing.
7. If you cannot complete all surveys in the selected house (because no respondent of the type you are looking for lives there), visit the closest house until all surveys have been completed. Continue doing this procedure until you complete all survey types as instructed.

8. After completing all survey types, select another starting household (or section and then household) at random if there is more than one interview location in the community or continue to the next selected interview location.

Remember: For each questionnaire, randomly select a starting household and then go to the closest house until all surveys are complete.

Manager's Checklist
(Pre-survey, During, and Post-survey)

The following is a checklist for program managers:

1. Review the data collection plan with each interviewer and supervisor.
2. Indicate the minimum number of interviews to be completed in one day.
3. During day 1 you can let data collectors work in pairs if you think this will increase their confidence.
4. Provide the technical and administrative support required by each interviewer (transport, lunch, etc.).
5. At the end of each day always review the questionnaires of each interviewer to assure that they have been correctly filled out and are complete. Check for any missing information or responses, and missing pages.
6. Make necessary corrections to questionnaire and inform the interviewer of problems found. If information is missing, the interviewer should revisit the house to complete the questionnaire before going to another community.
7. Confirm that all questionnaires have been filled in for each supervision area and that no pages are missing. If your LQAS sample size is 19 then you should have 19 completed questionnaires.
8. Organize the questionnaires by LQAS number (for example – from 1 to 19), according to the supervision area. For five supervision areas, for example, you would organize the questionnaires as follows if you had two types of respondents (young people 15-49 and women with children less than 1 year old for example):

   Folder 1: Respondent Type A, Supervision Area 1: 01 to 19
   Folder 2: Respondent Type A, Supervision Area 2: 01 to 19
   Folder 3: Respondent Type A, Supervision Area 3: 01 to 19
   Folder 4: Respondent Type A, Supervision Area 4: 01 to 19
   Folder 5: Respondent Type A, Supervision Area 5: 01 to 19

   Folder 6: Respondent Type B, Supervision Area 1: 01 to 19
   Folder 7: Respondent Type B, Supervision Area 2: 01 to 19
   Folder 8: Respondent Type B, Supervision Area 3: 01 to 19
   Folder 9: Respondent Type B, Supervision Area 4: 01 to 19
   Folder 10: Respondent Type B, Supervision Area 5: 01 to 19

   Etc…

Tabulation Checklist

As you tabulate your questionnaire, use the following checklist.

Before you begin:

1. Be sure the questionnaires you are about to tabulate match the type of tabulation table you have (right age, sex, etc.)
2. Confirm that questionnaires are in the correct order: 01 – 19, and confirm they have the correct number of pages.

During tabulation:

1. Work in threes.
2. The first person reads the correct answer on the tabulation sheet.
3. The second person looks at the answer on the questionnaire, determines if the answer is a "1" correct or a "0" incorrect. Mark an "S" for intentionally skipped questions that cannot be judged as either correct or incorrect, and an "X" for questions that should have responses but the responses are missing. An "X" should be taken out of the denominator.

   An "S" should only be marked if the person should be taken out of the denominator. In such cases it may not be possible to assess information for the SA but a program-wide estimate might still be possible. For example, if the question concerns whether the respondent was sick but the respondent has not been sick, then all the questions about the sickness would be marked as "S" since they are irrelevant for this respondent.

   However, in most cases a skipped question is equivalent to an automatic incorrect and should be coded as "0." For example, if a respondent says he/she has never heard of HIV, all subsequent questions about how to prevent HIV would be automatically incorrect. Similarly, if a respondent did not have any pre-natal visits during her last pregnancy, then all questions about counseling and testing would be coded as "0".
4. The first person records the answer on the tabulation sheet.
5. The third person confirms that the second person correctly determined if the answer should be coded "1" or "0" or "S" or "X" and that the first person recorded it properly.

After completing each column (all responses from one respondent):

1. Check that all the marks are in the same column; there should be no marks in the column to the right of the column just completed.
2. Check that there are no blank cells in the column just completed.
   - Be sure that no cells are blank. For any blank cell review the questionnaire to see if it should be coded a 0, 1, S, or X.

- Almost all responses should be a 0 or 1.
- If the cell has an "S," then check to see that it satisfies this criterion: The respondent was skipped because the question should not be asked of her/him because she/he cannot be included in the denominator. In a way, this means they are not part of the universe being assessed.

  EXAMPLE 1: Some questions are asked if the person was sick in the last 2 weeks. People who were not sick are coded "S" because the question cannot be asked of them.

  EXAMPLE 2: Some questions are not asked because the questions are automatically INCORRECT or 0. If a person is asked if she has ever heard of HIV/AIDS, and responds "No," that question is coded a 0 since it is not the desired response—it is incorrect. Any following question that asks about how HIV is transmitted or prevented would be SKIPPED since they are automatically counted as INCORRECT (coded "0") since we know the person cannot know the correct response because she does not even know that HIV exists.
- If the cell has an "X," this means the respondent should have responded to the question but for some reason no response was recorded. This could be because the interviewer forgot to do this. **Sometimes an interviewer circles several responses when they should have only circled one of them. These responses are also coded as "X" since there is no clear response. Also, if you cannot decipher the response written on a questionnaire, then "X" is an appropriate code. All "Xs" are excluded from the denominator in any calculation.**

3. Ask a manager or supervisor to check your tabulation sheet after you have completed the first column.

After completing a tabulation sheet:

1. Enter the total number correct in the appropriate column.
2. Enter the total sample size in the appropriate column.
3. Look at all questions where the sample is less than 19 and confirm the reason:
4. All questions should have a "0", "1", "S", or "X." If this is not the case, find out why, so you can make an appropriate entry in the space provided.

Elements of Survey Reports

Elements of Baseline and Monitoring Reports

	Content Baseline Report	Content Monitoring Report	Maximum Pages (13, excluding other appendices)
1	Summary	Summary	1
2	Program Overview (Locations, Objectives, Proposed Activities, Beneficiaries, etc.)	Program Overview (Locations, Objectives, Completed Activities, Beneficiaries, etc.)	1
3	Purpose of Baseline Survey and Methodology (Establish Priorities and Targets)	Purpose of Monitoring Survey and Methodology (Assess Progress Toward Targets)	1
4	Main Findings: Priorities by Supervision Area and for the Program as a Whole	Main Findings: Accomplishments, Priorities by Supervision Area and for the Program as a Whole	5
5	Action Plans and Goals/Coverage Targets for Key Indicators	Action Plans and Goals/Coverage Results (and Revised Targets) for Key Indicators	2
6	Conclusions and Next Steps	Conclusions and Recommendations	2
7	Appendix (Photocopy of Summary Tabulation Tables)	Appendix (Photocopy of Summary Tabulation Tables)	1

TOOLS AND REFERENCES

TOOLS AND REFERENCES

Some Resources

This document provides references for professionals who want more details about specific topics. This is not an exhaustive list of all references, but it contains references that are useful in designing and conducting surveys—especially related to HIV and reproductive health.

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

1. Rapid Health Survey Manual: http://robbresearch.pbworks.com/w/page/23229266/Rapid-Health-Surveys:-Principles-and-Sampling-Design-Handbook
2. The Maternal and Child Health Integrated Program (MCHIP) Non-Governmental Organization (NGO) strengthening site: www.MCHIPNGO.net or www.childsurvival.com
3. The CORE Group: www.coregroup.org
4. The Demographic Health Survey (DHS) site: www.measuredhs.com
5. MEASURE Evaluation: www.cpc.unc.edu/measure
6. The Research Methods Knowledge Base: http://www.socialresearchmethods.net/kb
7. The Flexible Fund: www.flexfund.org
8. UCLA Department of Epidemiology, Rapid Survey Course: http://www.ph.ucla.edu/epi/rapidsurveys/RScourse/RSrapidsurveys.html

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.

General Surveys
(With Focus on Knowledge, Practice and Coverage Surveys Useful to NGOs)

Useful Documents	Key points About Documents
Rapid Health Survey Manual http://robbresearch.pbworks.com/w/page/23229266/Rapid-Health-Surveys:-Principles-and-Sampling-Design-Handbook A pdf version of the manual is available here: http://www.mchipngo.net/lib/components/documents/MandE/PHI_RapidHealth.pdf	• This manual is a "one-stop" resource for information on conducting rapid health surveys. It is not a training manual but provides information on cluster sampling and LQAS with clear examples of how to conduct surveys at the community level.
Knowledge Practice and Coverage Resources (KPC) Many resources available through this link: http://www.mchipngo.net/controllers/link.cfc?method=tools_mande The "Field Guide" is available at: http://www.mchipngo.net/lib/components/documents/KPC/FieldGuide_Sept03.pdf The Methodology and Sampling Issues for KPC Guide is available at: http://www.mchipngo.net/lib/components/documents/KPC/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://coregroup.org/storage/documents/Workingpapers/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
KPC Trainer of Survey Trainers (TOAST)[1] http://coregroup.org/our-technical-work/working-groups/monitoring-and-evaluation	• 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 are available at this page: http://coregroup.org/our-technical-work/working-groups/monitoring-and-evaluation	• 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://coregroup.org/storage/documents/Workingpapers/LQAS_Protocol_for_Parallel_Sampling.pdf	• Practical instructions for parallel sampling with LQAS. Developed from a technical advisory meeting on the subject.

Reproductive Health: The Flex-Fund

Useful Documents	Key Points About Documents
Program Design Monitoring and Evaluation (PDME) Curriculum http://www.flexfund.org/resources/training/pdme.cfm For indicators go here: http://www.flexfund.org/resources/grantee_tools/tab_plan.cfm For questionnaires go here: http://www.flexfund.org/resources/grantee_tools/survey_quest.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.

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

HIV/AIDS Resources-Measure/DHS

Useful Documents	Key Points About Documents
AIDS Indicator Survey (AIS) http://www.measuredhs.com/aboutsurveys/ais/start.cfm For indicators go here: http://www.measuredhs.com/hivdata/ind_tbl.cfm For questionnaires go here: http://www.measuredhs.com/aboutsurveys/ais/questionnaires.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.
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.

Demographic and Health Surveys

Useful Documents	Key Points About Documents
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.

Useful Documents	Key Points About Documents
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

MEASURE Evaluation has developed a variety of monitoring and evaluation tools and guides. The following are important examples.

Other DHS/MEASURE Evaluation

Useful Documents	Key Points About Documents
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.

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.

UCLA Department of Epidemiology Rapid Survey Course[5]

Useful Documents	Key Points About Documents
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.

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