KPC 2000+

Knowledge, Practices, and Coverage Survey 2000+

FIELD GUIDE

Developed by Donna Espeut,
The Child Survival Technical Support Project

In July 2001, the KPC Revision Task Force of the CORE Monitoring and Evaluation Working Group met to review the April 2001 draft of the *KPC2000+ Field Guide*. The current version of the guide (August 2001) incorporates recommendations from the Task Force.

In its present form, the *KPC2000+ Field Guide* familiarizes the reader with important issues and concepts related to KPC surveys.

CORE and CSTS encourage you to provide additional input in further developing this guide. Please submit your comments, questions, or suggestions to CSTS+. The contact information is provided below.

The Child Survival Technical Support
Project+
(ATTN: Donna Espeut)
11785 Beltsville Drive
Calverton, Maryland 20705
U.S.A.

TEL: (301) 572-0200
FAX: (301) 572-0983
E-MAIL: csts@orcmacro.com

Special thanks to Counterpart International, Inc. and consultants Tom Davis and Julie Mobley for field-testing preliminary components of the field guide in January 2001.

TABLE OF CONTENTS

LIST OF ACRONYMS

ARI	acute respiratory infection
BCC	behavior change communication
BHR/PVC	Bureau of Humanitarian Response and Private and Voluntary Cooperation
CATCH	Core Assessment Tool on Child Health
CS	child survival
CSSP	Child Survival Support Program
CSTS	Child Survival Technical Support Project
DE	design effect
DHS	Demographic and Health Surveys
DIP	Detailed Implementation Plan
HIS	Health Information System
HIV/AIDS	Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome
IMCI	Integrated Management of Childhood Illnesses
KPC	Knowledge, Practices, and Coverage survey
LQAS	Lot Quality Assurance Sampling
M&E	monitoring and evaluation
MEWG	CORE Monitoring and Evaluation Working Group
MICS	Multiple Indicator Cluster Survey
MOH	Ministry of Health
NGO	nongovernmental organization
ORS	oral rehydration salts
ORT	oral rehydration therapy
PPS	probability proportional to size
PVO	private voluntary organization
SPA	Service Provision Assessments
SRS	simple random sampling
STI	sexually transmitted infection
USAID	United States Agency for International Development

INTRODUCTION TO GUIDE

Why Was This Guide Written?

The private voluntary organization (PVO) Child Survival Support Project (CSSP) of The Johns Hopkins University initially developed a document known as the *Survey Trainer's Guide for PVO Child Survival Project Rapid Knowledge, Practice, and Coverage (KPC) Surveys*. The guide aimed to help standardize field implementation of KPC surveys. Its target audience was individuals who attended one of the many Training of Survey Trainers Workshops organized by PVO CSSP.

The present guide, the *KPC2000+ Field Guide*, is an updated version of the *Survey Trainer's Guide*. The new guide corresponds to the October 2000 version of the KPC, which is now called the *KPC2000+*. The *KPC2000+* is a revision of the original questionnaire developed by PVO CSSP, and is a product of the Child Survival Technical Support Project (CSTS) and the CORE Monitoring and Evaluation Working Group (MEWG).

Like the original KPC manual, the *KPC2000+ Field Guide* aims to assist projects in planning, conducting, and analyzing a KPC survey. The present guide contains expanded sections on sampling options for KPC surveys, KPC data analysis, and the use of KPC data for health decisionmaking. It also stresses the importance of incorporating qualitative research, partnership-building, and capacity development into the KPC process.

For Whom Was This Guide Written?

The *KPC2000+ Field Guide* was written for persons who will be conducting KPC surveys but have not yet had the opportunity to attend a KPC training workshop. An effort has been made to present concepts in a simple and easy-to-understand manner. However, this guide is not intended to take the place of formal instruction in survey research. Individuals who do not have a basic understanding of survey research and program monitoring and evaluation are encouraged to refer to other resources (such as the ones listed at the end of each section) and/or seek the assistance of individuals with such experience.

What Does This Guide Contain?

As mentioned above, the guide is not designed to be a stand-alone resource when conducting a KPC survey. It does, however, provide a comprehensive overview of the KPC process. More specifically, it covers the following:

- Purpose of a KPC survey and its role in project monitoring and evaluation
- *KPC2000+* tools
- Phases of the KPC process
- Useful research materials produced by other agencies and organizations.

PURPOSE AND ROLE OF A KPC SURVEY

History of the Rapid KPC Survey

PVOs have played a major role in improving the health outcomes of women and children worldwide. However, a shortage of staff with monitoring and evaluation (M&E) training has made it difficult to document progress. In response to the need for a rapid, easy-to-use way of assessing progress, the United States Agency for International Development (USAID) asked the PVO Child Survival Support Program (CSSP) to develop an assessment tool. The PVO CSSP, the original technical assistance project for the PVO child survival (CS) community, consulted with PVO staff and designed the Rapid Knowledge, Practices, and Coverage (KPC) Survey. The KPC is a management tool that yields a concise and manageable set of indicators to monitor and estimate the results of PVO CS activities. In addition, survey implementation is intended to promote local participation in identifying health priorities and in monitoring community health status.

CS projects generally aim to improve the health and nutrition outcomes of children under the age of five years. However, the KPC survey usually targets mothers of children under the age of two years. Reasons why the KPC has focused on children under age two are listed below.

- Among children under 5, children under 2 experience the highest health risks.
- Budget and human resource constraints warrant limiting the age range of children who are surveyed to those under 2.
- Projects are given a short period of time to establish interventions and assess impact. As a result, some PVOs choose to monitor and estimate program effects based upon beneficiaries who are under 2. If children under 5 are included in assessments, the effects of a program may be diluted by the experiences of older kids who were not program beneficiaries.

In the past, it was common practice to only include mothers of children under 2 in a KPC survey. As a result, children with non-maternal caregivers (for example, grandmothers or older siblings) were excluded from the survey. Children whose mothers are absent from the home, leaving the child in the care of other individuals, might have greater chances of getting ill or dying compared with children who are cared for by the biological mother. The HIV/AIDS epidemic is changing the social and economic realities in many countries, and PVOs are starting to re-consider who should be included in a KPC survey. As a result, more KPC surveys are including children under 2 and their primary caregivers, regardless of whether the caregiver is the biological mother or someone else. This more accurately reflects the population of target beneficiaries.

Traditionally, mothers have been selected for the survey using a 30-cluster sampling methodology. This method is an efficient way of obtaining coverage estimates for an entire program area.

Today, projects are trying to maximize the amount and type of information collected in a KPC survey by experimenting with the instrument, the methodology, and the analysis. Below are some examples.

Modify the instrument

- Develop separate questionnaires for emerging targets of CS projects such as non-maternal caregivers (grandmothers, siblings) and adolescents
- Modify the structure of the questionnaire to collect information on children of different age groups (0–11 months, 12–23 months) and on children who experienced an illness in the past 2 weeks

Modify the methodology

- Use parallel sampling techniques to collect information from groups of interest other than mothers, such as fathers, adolescents, or other women of reproductive age
- Use Lot Quality Assurance Sampling (LQAS) to make community-level assessments and allocate project resources in a more targeted manner
- Collect information on more than one child, not just the youngest child under 2

Modify the analysis

- Explore differentials between subgroups within the sample (for example, compare girls to boys, maternal caregivers to non-maternal caregivers, young mothers to older mothers)

Projects have not experimented with multi-level analysis, but there is also the potential to link KPC data with facility-level and community-level data from other sources.

The KPC used to be required of all USAID-funded CS projects as part of their baseline and final assessments. Although USAID no longer requires KPC surveys, CS projects still rely on the survey for useful beneficiary-level information.

Recent revisions to the KPC

In recent years, PVOs expressed a desire to revise the original KPC to include topics such as anthropometry, malaria, and HIV/STIs. CSTS and the CORE MEWG were responsible for revising the questionnaire. The current version, the *KPC2000+*, consists of the following:

- The Rapid CATCH (Core Assessment Tool on Child Health)
- Fifteen modules that correspond to key CS technical interventions.

The Rapid CATCH questions are linked to a concise set of indicators that reflect current international standards in child health and survival. In adapting the KPC to reflect project

activities and objectives, PVOs are encouraged to add questions from relevant modules to the *Rapid CATCH*. *The section of this guide entitled "KPC2000+ Tools" further describes the Rapid CATCH and the modules.*

The Role of KPC Surveys in a Child Survival Project

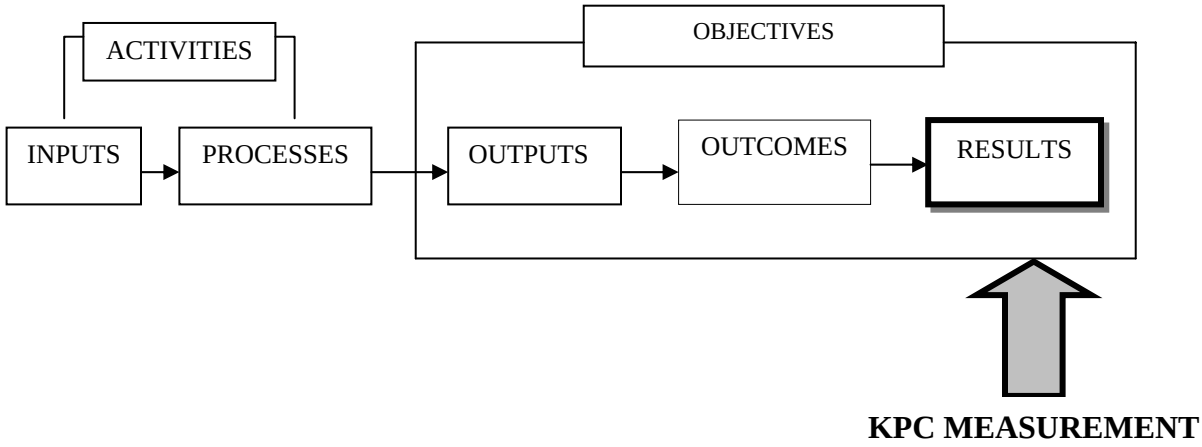
The KPC within a Results Framework

The ultimate goals of a CS project are to reduce under-5 morbidity and mortality. Individual projects usually operate with limited resources and within a limited time frame, making it difficult to achieve these goals. Projects therefore focus on results *s*uch as the following:

- Improved maternal nutritional status
- Improved child nutritional status
- Timely and complete immunization of young children
- Appropriate case management of common childhood illnesses
- Widespread practice of behaviors that reduce the risk of common childhood illnesses
- Antenatal care coverage
- Safe deliveries
- Postpartum contact with a health provider
- Adequate child spacing
- Prevention and early detection of HIV/STIs
- Environmental conditions that are conducive to disease prevention.

Project *objectives* and *activities* relate to these results, and projects select *indicators* to track their achievement of objectives. In addition to end results such as the ones listed above, there are intermediate results, also known as *outcomes*. Outcomes can occur at various levels but should lead to project results and ultimately, CS goals.

Having a clearly defined framework of how objectives, activities, and indicators link to outcomes and results is critical to good program management. Below is a Results Framework for project monitoring and evaluation.



Children and their caregivers are the ultimate beneficiaries of a CS project. However, projects often operate at many different levels to achieve their CS objectives. Project activities may relate to one or more of the following:

THE PVO
- PVO
- Health unit of the PVO
- CS project

LOCAL PARTNERS
- Local nongovernmental organizations (NGOs)
- Private partners
- Ministry of Health (MOH)
- Districts/municipalities
- Health facilities

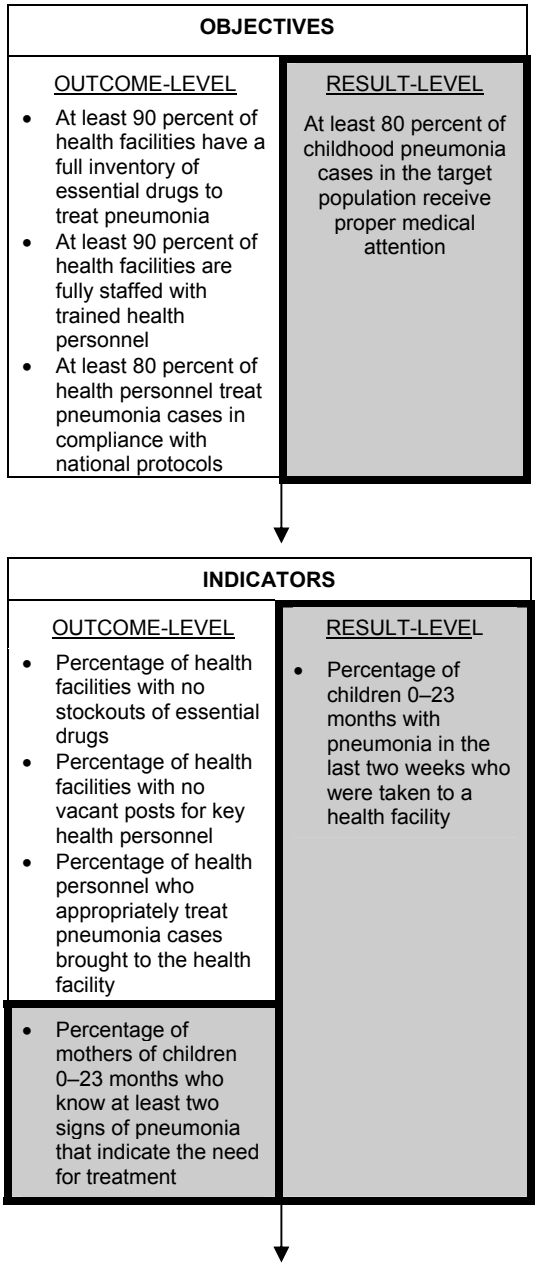
THE COMMUNITY/INDIVIDUAL
- Communities
- Community-based organizations
- Community health workers
- Private providers
- Households
- Mothers/caregivers
- Children under 5.

FYI—For Your Information

CSTS is working on a Capacity Tool Bank that depicts the various levels of capacity strengthening, as well as examples of objectives, indicators, tools, and activities at each of those levels. Visit the CSTS Web site at www.childsurvival.com/tools/tool_start.cfm to access the Capacity Tool Bank. Click on each box to see references at that level.

KPC DATA

The table on the following page provides a concrete example of the Results Framework. As seen by the highlighted boxes, a KPC generally relates to results at the individual (caregiver or infant/child) level. Please note that it is also important to measure outcomes. Some projects operate at the health-systems level (for example, improving health-worker performance or strengthening the district health system) rather than focus on changing individual behaviors. Project indicators should always reflect project objectives and activities. The measurement of outcomes or results at levels beyond the individual will require tools and methods other than a KPC survey.



ACTIVITIES	OUTPUTS	OUTCOMES	RESULT	GOAL
• Facility-based health worker training in IMCI or pneumonia case management • IEC activities targeting caregivers	• Number of facility-based health workers trained in IMCI/pneumonia case management • Number of radio spots promoting recognition of pneumonia danger signs and prompt treatment • Number of educational sessions (with mothers groups, community health workers, etc.) held in target communities	• Availability of essential drugs at health facilities • Availability of trained health personnel at health clinics and health posts • Appropriate management of pneumonia cases in health facilities • Caregiver recognition of pneumonia danger signs indicating the need for treatment	Timely and appropriate pneumonia case management	Reduced pneumonia-related mortality in children under 5

When to Conduct a KPC Survey

A KPC survey can be implemented at the beginning, during, or at the end of a project cycle. The following page depicts a timeline for M&E activities after a PVO receives donor funding. In the past, CS projects have been funded for a 4-year cycle. USAID now funds 5-year projects.

At the beginning: When conducted at the beginning of a project, a KPC survey can assist projects in identifying and prioritizing problems within the project's catchment population.

During: Although the KPC is a rapid assessment technique, it is not very practical to implement during the life of a project. However, parts of a KPC—for example, 10–20 questions that relate to sentinel indicators of CS performance—can be used with a variety of methods and techniques to monitor activities.

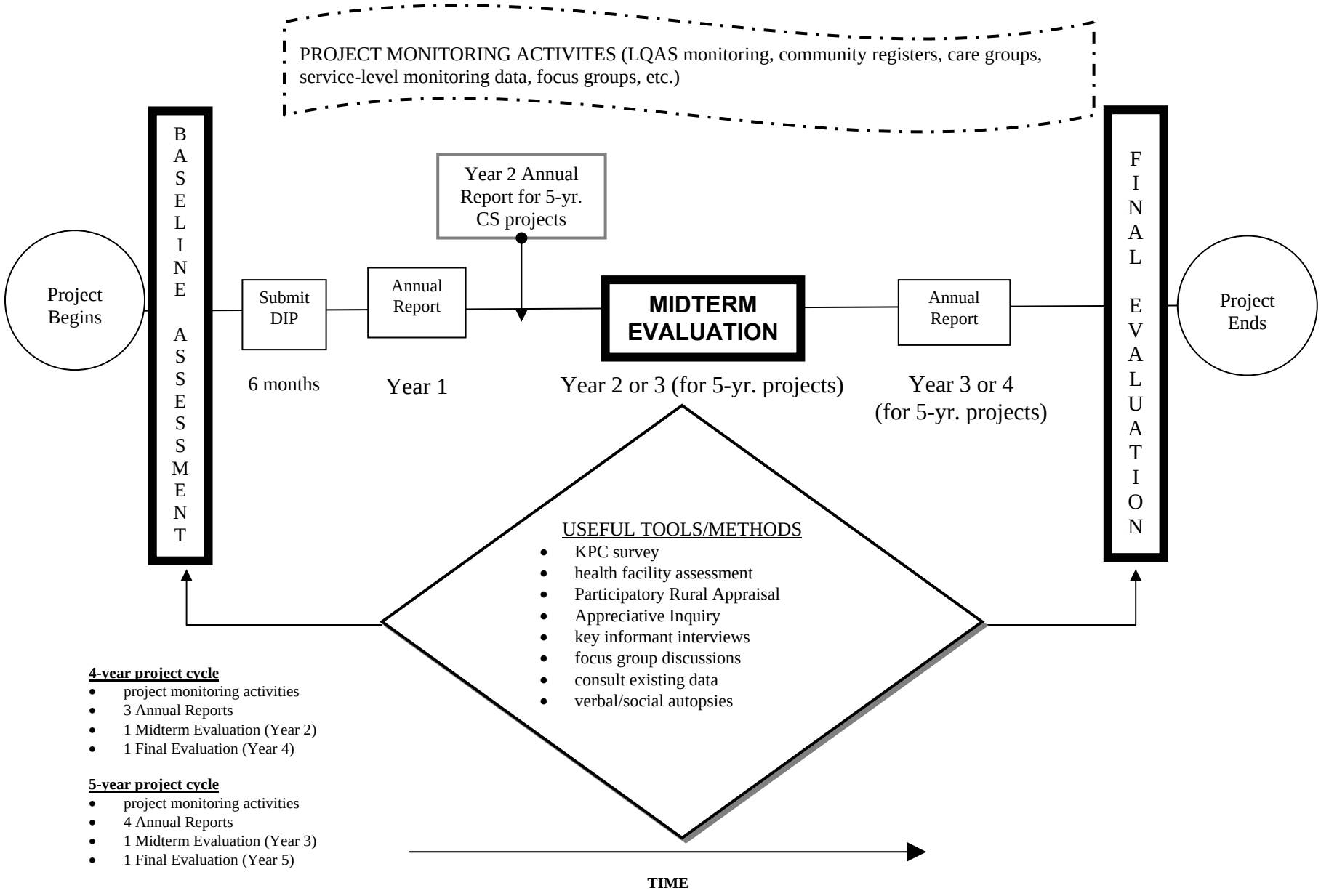
At the end: KPC surveys are often conducted at the end of a project to assess whether a project achieved its initial objectives.

As shown in the timeline, a KPC survey is just one of many tools that can be used for M&E. Although projects should aim to be as comprehensive as possible in their assessments, they are also encouraged to collect information that is programmatically useful. It is not necessary to use every possible tool or method available, nor is it necessary to use every single KPC question. Instead, projects are encouraged to minimize the amount of time and resources invested in data collection and maximize the amount spent using and disseminating information.

REMEMBER:

The KPC is just one of many tools that can be used to identify problems and track progress. There are different types of tools that apply to the different levels of capacity strengthening. In addition, there are various methods of information gathering. Everything cannot be easily quantified. It is often necessary to use a combination of qualitative and quantitative methods to better understand the factors and processes that impact child health and survival.

GENERAL TIMELINE FOR CS M&E ACTIVITIES



What a KPC Can and Cannot Do

A KPC survey can yield highly useful information on children and their caregivers and, to a lesser degree, households and communities. Projects can use KPC surveys to build consensus with local partners and develop local capacity for information gathering, analysis, and the use of information for decisionmaking.

In addition to knowing its strengths, it is important to recognize the limitations of a KPC survey. Below are two short lists of what a KPC can and cannot do.

WHAT CAN A KPC DO?

- ✓ Estimate individual-level results
- ✓ Assist projects to identify and prioritize problems that exist within their project area (conducting a baseline KPC)
- ✓ Aid projects in determining whether objectives have been achieved (conducting a KPC at the end of a project)
- ✓ Develop local capacity to collect, analyze, and use information for decisionmaking
- ✓ Help build consensus between projects, local partners, and stakeholders
- ✓ Enable projects to track their progress in achieving CS objectives (e.g., using a small number of KPC questions linked to benchmark indicators for project monitoring)

WHAT CAN'T A KPC DO?

- **x** Address other levels of capacity strengthening, namely the PVO, its local partners, and the community at large
- **x** Document process
- **x** Document the achievement of <u>goals</u> such as a reduction in under-5 mortality
- **x** Measure impact, unless major modifications are made to the following:
  - The project design (e.g., using non-intervention ("control") groups or communities)
  - The survey design (e.g., increasing sample sizes of baseline and final KPC surveys to be able to measure changes in indicators over time)

The remainder of this manual will guide you through each of the steps in designing, conducting, and analyzing a KPC survey.

KPC2000+ TOOLS

A major revision to the original KPC questionnaire is the creation of 15 separate modules. There is one module for each of the key child survival interventions. Each module contains the following:

- Interviewer instructions
- Examples of questions/themes that can be explored using qualitative research
- Suggested quantitative research questions
- Basic tabulation plan with key indicators.

REMEMBER:

When reviewing the modules, remember to read the footnotes, which contain useful information when adapting the survey. For example, cross-referencing between modules draws attention to questions located in different modules that are related to the same topic area.

The newest feature of the KPC is *Rapid CATCH*. The tool contains 26 questions from the *KPC2000+* modules and yields 13 key child health indicators. Projects are given two options in terms of immunization indicators. One is a card-based indicator for full immunization before the first birthday, and the other is a measles indicator based solely on the mother's recall. *Rapid CATCH* provides a snapshot of the target population in terms of child health. It represents the bare minimum in terms of issues that CS projects should consider in their assessments. *Rapid Catch* is not context-specific, however. Therefore, PVOs should add selected questions from the modules so that the KPC reflects their program objectives, activities, and the contexts in which they are working.

It is likely that no two KPC questionnaires will look alike, although *Rapid CATCH* questions should be embedded within each questionnaire—regardless of project intervention mix. The CORE MEWG strongly encourages all CS projects to report the *CATCH* indicators, which provide critical information on life-saving household behaviors and care-seeking patterns that affect the health and survival of children worldwide. A statement from the MEWG appears at the beginning of *Rapid CATCH*.

REMEMBER:

Rapid CATCH serves as a starting point for projects when developing a KPC survey. It is simply a subset of important questions from the 15 modules. The modules provide additional questions that allow each project to adapt the survey to fit a specific context. In the modules, the highlighted questions are the *Rapid CATCH* questions.

REMEMBER:

There is a Tabulation Plan at the end of *Rapid CATCH* that provides guidance in tabulating priority child health indicators.

The remainder of this section on tools briefly outlines the content of *Rapid CATCH* and the 15 modules that make up the *KPC2000+*.

Topics covered in Rapid CATCH (26 questions)

Topic	Question Number(s)
Interview date	1
Respondent age	2
Child spacing/household under-5 density	3–5
Child anthropometry	6–7
Maternal/newborn care	8–10
Breastfeeding/nutrition	11–13
Child immunization	14–16
Malaria Prevention	17–19
IMCI	20–23
HIV/AIDS	24–25
Handwashing	26

Topic Covered in the 15 Modules

Module 1A: Water and Sanitation (10 questions):

- Month of year
- Source of drinking water
- Toilet facility
- Waste disposal
- Handwashing

Module 1B: Respondent Background Information (7 questions):

- Years of schooling
- Languages/dialects spoken
- Household structure
- Gainful employment
- Caregiver when mother is away from home

Module 2: Breastfeeding and Infant/Child Nutrition (11 questions):

- Ever and current breastfeeding
- Initiation of breastfeeding
- Provision of colostrum and prelacteal feeds
- Duration of breastfeeding
- Food/liquid consumption in past 24 hours
- Salt iodization
- Vitamin A supplementation

Module 3: Growth Monitoring and Maternal/Child Anthropometry (12 questions):

- Weighing of infant at birth
- Growth monitoring
- Deworming
- Child height/weight
- Maternal arm circumference

Module 4A: Childhood Immunization (6 questions):

- Vitamin A supplementation
- Immunization card possession
- Immunizations received

Module 4B: Sick Child (4 questions):

- Caregiver knowledge of child danger signs
- Illnesses in the 2 weeks before the survey

Module 4C: Diarrhea (14 questions):

- Diarrhea treatment/management
- Sequence of care-seeking
- Knowledge of ORS preparation
- Household hand-washing facility
- Hand-washing behavior

Module 4D: Acute Respiratory Infections (ARI) (10 questions):

- ARI treatment
- Sequence of care-seeking

Module 4E: Malaria (24 questions):

- Treatment/care-seeking for fever
- Causes of malaria
- Malaria prophylaxis during pregnancy
- Bednet use, maintenance, and quality

Module 5A: Prenatal Care (13 questions):

- Frequency/nature of prenatal check-ups
- Tetanus toxoid immunization
- Maternal health card possession
- Access to nearest health facility
- Decisionmaking for facility care-seeking
- Knowledge of pregnancy danger signs

Module 5B: Delivery and Immediate Newborn Care (8 questions):

- Place of delivery
- Delivery assistance
- Delivery practices

Module 5C: Postpartum Care (11 questions):

- Postpartum check-ups
- Provider of postpartum care
- Knowledge of postpartum danger signs in the mother
- Newborn danger signs
- Content of postpartum care
- Maternal Vitamin A supplementation

Module 6: Child Spacing (8 questions):

- Household under-5 density
- Length of previous birth interval
- Knowledge of sources of child spacing methods
- Desire for more children
- Current contraceptive use
- Postpartum information on child spacing

Module 7: HIV/AIDS (57 questions):

- Knowledge of HIV/AIDS risk factors/modes of transmission
- Risk and risk reduction
- STIs
- HIV screening
- Stigma
- Sources of care and support
- Orphans/foster children

Module 8: Health Contacts and Sources of Information (3 questions):

- Contact with different health providers
- Frequency of contacts
- Sources of information on health/nutrition
- Exposure to health messages by source

PHASES OF A KPC SURVEY

In this guide, the KPC process has been divided into the following three phases:

- **The Pre-implementation Phase,** which involves activities such as meeting with project stakeholders and local experts, assessing data needs, developing a questionnaire, designing a sampling strategy, and training supervisors and interviewers
- **The Field Implementation Phase,** which involves the actual collection of data in selected communities
- **The Post-implementation Phase,** which involves tabulating and analyzing the data, disseminating findings, and using the data for health decisionmaking.

KPC SURVEY TIMELINE

(Estimated total duration=28 days)

PRE-IMPLEMENTATION PHASE (18 Days)

Assess Information Needs and Incorporate Stakeholders and Local Experts

- Review existing project documents (e.g., proposal, DIP, annual reports)
- Consult existing data sources
- Meet/collaborate with community members and other project stakeholders
- Invite MOH statisticians and/or individuals from local universities/institutions who have expertise in survey research

Make Administrative and Logistical Preparations

- Form KPC Coordinating Team comprised of project stakeholders and local experts
- Determine survey dates/timeline
- Develop logistics plan for all phases of survey

Determine What/From Whom/How Information Will Be Collected

- Identify survey indicators and develop analysis plan
- Conduct formative research to identify local information needs, terms/concepts to be included in survey
- Identify survey targets (e.g., mothers, non-maternal caregivers, men of reproductive age, adolescents)
- Design sampling strategy
- Select sample areas (clusters, lots, etc.)
- Adapt/translate/back-translate questionnaire(s)
- Pretest questionnaires; revise as necessary based on pretest

Maximize Quality of Data Collected

- Develop/adapt training materials for supervisors/interviewers
- Develop procedures for quality control during data collection/data entry
- Recruit and train field personnel (supervisors, interviewers, etc.)
- Field-test questionnaires with interviewers/supervisors; make final revisions, then photocopy final version of questionnaires

Data Management

- Develop data entry, error-checking, and data analysis programs
- Create hand tabulation tables
- Plan tabulation, analysis, and feedback sessions (post-data collection)

SURVEY IMPLEMENTATION PHASE (4 Days)

- Conduct interviews
- "Clean" questionnaires in the field before tabulation and analysis
- Enter data electronically as questionnaires are reviewed and submitted by field supervisors
- Refine analysis programs (for computer analysis only) and hand tabulation tables
- Finalize logistics for tabulation/analysis workshop(s)

POST-IMPLEMENTATION PHASE (6 Days)

- Complete data entry/data cleaning
- Tabulate/analyze survey data
- Identify and prioritize problems
- Draft survey report
- Hold community feedback session(s)
- Share preliminary findings with mission, MOH, and other interested parties
- Develop action plans addressing problems identified by survey
- Conduct follow-up qualitative research or special surveys, if necessary
- Design ways of displaying KPC findings

General Timeline for Conducting a KPC Survey

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

INCORPORATE LOCAL STAKEHOLDERS AND EXPERTS THROUGHOUT THE KPC PROCESS.

PRE-IMPLEMENTATION PHASE

Although the KPC is designed to be a rapid assessment, there are many steps in preparing for a KPC survey. Projects will probably require 2–3 weeks of pre-implementation activities.

During the Pre-implementation Phase, a project will prepare the following written documents:

- KPC questionnaire that reflects the local context
- Supervisor/interviewer training curriculum
- Data management guidelines
- Tabulation and analysis plan
- Sampling frame and maps for the study.

By the end of this phase, interviewers and supervisors should be trained and recruited, and the project should have all official clearances to conduct the survey. Although dissemination activities will not take place until after the data have been collected and analyzed, it is not too soon to start planning dissemination activities with local partners and stakeholders.

The checklist on the following page outlines key tasks that should be accomplished before the beginning of data collection.

✓ PRE-IMPLEMENTATION CHECKLIST

Process

- ❑ Invite MOH and local government staff to participate in survey process
- ❑ Consult with local partners/stakeholders to discuss nature and level of their involvement in KPC activities

Questionnaire

- ❑ Review project documents
- ❑ Assess information needs of the PVO and local partners/stakeholders
- ❑ Consult with local experts (e.g., individuals at universities with survey or statistics expertise, MOH statistician) and officials (local government staff)
- ❑ Refer to existing data sources
- ❑ Identify project indicators best measured by a KPC
- ❑ Design survey questionnaire
- ❑ Translate and back-translate the questionnaire
- ❑ Pretest questionnaire and revise accordingly
- ❑ Duplicate (photocopy) questionnaires

Methods

- ❑ Create maps of project area
- ❑ Select geographic area(s) to be surveyed
- ❑ Determine survey targets and units of analysis
- ❑ Choose appropriate sampling methodology
- ❑ Determine sample size and identify sampling points
- ❑ Reach consensus on quality-control procedures and develop data management guidelines
- ❑ Develop tabulation and analysis plan based upon study indicators and context

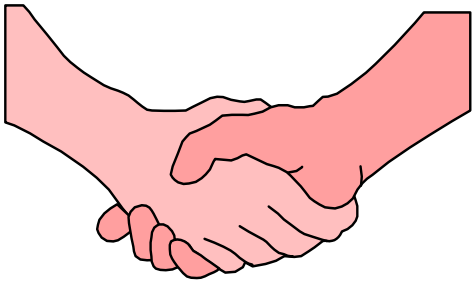
Personnel

- ❑ Assemble coordinating team of project staff and local partners; identify Survey Coordinator to oversee entire KPC process
- ❑ Develop criteria for selection of supervisors and interviewers
- ❑ Develop or adapt training materials
- ❑ Recruit and train supervisors and interviewers
- ❑ Select additional personnel to support all aspects of survey
- ❑ Make transportation, food, and lodging arrangements for all field personnel, if necessary

Resources and Logistics

- ❑ Determine survey dates
- ❑ Develop survey timeline and budget
- ❑ Make logistical preparations for survey training, data collection, tabulation, and dissemination activities

1. MAKE THE PROCESS PARTICIPATORY



The KPC was designed to be participatory. It can help build consensus between CS projects and their local partners. It can also be used to develop local capacity to gather, analyze, and use information for decisionmaking. PVOs should view the KPC as an opportunity to strengthen capacity both within and outside of the project.

Collaboration with local agencies and organizations is useful at all stages of the project cycle. It is particularly important for PVOs who are new to a geographical region and want to use a KPC to obtain baseline information. Engaging individuals from established and respected local organizations can help diffuse community suspicions about a new project. Religious or community leaders can play an integral role in the process. It is also a good idea to invite local government staff, MOH statisticians, and/or individuals from local universities or institutes who have expertise in survey research to participate in the survey.

As a first step, review the project's proposal. Identify the goals, objectives, and indicators. It is helpful to publicly display this information, which might help project staff and stakeholders connect with what the project is striving to achieve. As the project collects KPC and other data, it can explore ways of visually displaying this information so that everyone can track the achievement of objectives.

In addition to the above example, there are many other ways to foster ownership in a project. Below are some ways to make the KPC survey more participatory.

Ways to identify the needs and concerns of local partners/stakeholders:

- Key informant interviews with community leaders
- Attendance at community meetings to learn about the perceived needs of stakeholders
- Meetings with staff from local agencies and organizations (governmental and nongovernmental)

Roles that local partners/stakeholders can play in a KPC survey:

- *Survey coordination*—invite one or two individuals from partner organizations to join the KPC Coordinating Team
- *Questionnaire development*—consult with local partners/stakeholders to a) identify important issues the survey should address and b) identify local terminology/concepts that make the questionnaire more context-specific
- Designing the sampling strategy
- *Data collection*—recruit local partners to be interviewers or supervisors
- *Data tabulation and analysis*—invite local partners/stakeholders to a hand tabulation workshop, where they can use information from the completed questionnaires to calculate key indicators; train local partners and/or project staff to use Expanded Program on Immunization (EPI)-info for data entry and data analysis; work with project stakeholders to identify and prioritize health problems identified in the KPC survey
- *Data dissemination*—work with local partners to hold community feedback sessions
- *Data use*—engage local partners in project design or in the development of action plans. For example, invite them to the Detailed Implementation Plan (DIP)-writing workshop.

REMEMBER:

A major part of partnership building is *listening* to the concerns and priorities of local partners and stakeholders.

THE KPC COORDINATING TEAM

One of the first steps in the KPC process is to form a coordinating team. Members of the coordinating team should have the ability to train others, as well as the ability to organize and supervise complex activities. It is also important that persons who make up the coordinating team be available during the entire survey process.

In an effort to build consensus with local partners, one or two individuals from partner agencies or organizations can be invited to serve as members of the coordinating team. Projects will need to identify a person who can serve as the Survey Coordinator. Ideally, this will be someone local. However, there are instances when a project cannot identify a local person with KPC training or experience. In this situation, the project may choose to hire an outside consultant as Survey Coordinator. Even when a consultant is hired to help plan and oversee the study, it is helpful to identify staff members who can work

closely with him or her and develop the necessary skills to plan, carry out, and analyze similar surveys in the future.

There are many resources on participatory research and evaluation. Three such resources are listed below.

Participatory Program Evaluation Manual: Involving Program Stakeholders in the Evaluation Process (Aubel, 1999). **Available on CSTS Web site (www.childsurvival.com).**

Participatory Community Planning for Child Health: Implementation Guidelines. Arlington: BASICS (Bhattacharyya & Murray, 1999). **Available on BASICS Web site (www.basics.org).**

Qualitative Research for Improved Health Programs: A Guide to Manuals for Qualitative and Participatory Research on Child Health, Nutrition, and Reproductive Health (Winch et al., 2000). **Available on CSTS Web site (www.childsurvival.com).**

2. DEVELOP A LOGISTICS PLAN

Choosing Survey Dates—Some Things to Consider

- Are there holidays during the scheduled dates of the survey?
- What are weather conditions like at the scheduled time of the survey? If possible, avoid conducting the survey during monsoon season or at other times of the year when there are long periods of bad weather.
- What is the potential availability of respondents? For example, it might be very hard to conduct interviews with people during harvest season, when they are likely to be away from home for long periods of time.
- Will there be other project activities taking place at the same time that will create a major scheduling conflict for key persons involved in the survey?

Keep in mind that the time of year when you conduct the survey affects how representative survey findings will be of conditions in general. Disease prevalences, as well as food security and dietary practices, usually vary throughout the year.

REMEMBER

The KPC is a rapid assessment, but your project should plan to spend approximately 4 weeks engaged in KPC-related activities. Not all of this time will be full-time work. Many projects try to complete the survey in a much shorter period of time. Although a participatory approach takes a little more time, the extra time and effort spent to incorporate local partners/stakeholders, build local capacity, and properly train interviewers will result in high-quality information that is owned and used by all stakeholders. The benefits of this buy-in at the local level will extend far beyond the KPC survey.

Develop a budget

In developing a budget for the survey, it is recommended that you consider all resource requirements.

- Review requirements for the following:
  - ⇨ Personnel for conducting training sessions, interviews, tabulation, analysis, and

dissemination. This includes anticipated expenses in providing food and/or lodging for supervisors, interviewers, and other field personnel.

- ⇨ Supplies for copying questionnaires and hand-tabulation tables, training of supervisors and interviewers, conducting the survey, collecting anthropometric data (purchase/borrow scales and measuring boards), tabulation, analysis, and presentation of results.
- ⇨ Transportation for field-based practice, data collection, and feedback sessions (local and national).

- Develop a budget for survey costs (salaries, per diems, equipment, supplies, and room and board for project personnel and interviewers, rental of equipment and facilities, etc.)

For a permanent record of expenses that can be used when budgeting for future surveys, include the actual cost breakdown as an appendix in the KPC survey report.

Select personnel to support all phases of the survey

- The number of interviewers, supervisors, and days required to complete the survey will vary according to factors such as resources, weather conditions, and number of interviews. It might be useful to train extra staff in case of illness or some other emergency. Considerations include the availability of personnel and transport as well as travel factors. It is also important to consider the trade-offs between the number of interviewers, the length of data collection, and the quality of data. Having a small number of interviewers (and therefore, fewer survey teams) would increase the amount of time needed to complete all of the interviews, but this might increase the internal consistency of the information collected. Yet, increasing the number of people involved in data collection can promote greater ownership of the KPC process, results, and the project as a whole. Although a larger number of interviewers might also reduce the amount of time spent in the field, there is a possibility that there will be greater variation (less consistency) in the quality of interviews conducted.

The following table gives an indication of the minimum time requirements for supervisors and interviewers.

MINIMUM TIME REQUIREMENTS: SUPERVISORS AND INTERVIEWERS

Activities	Time Allocation (in days)	
	Supervisors	Interviewers
Training	4	3
Data Collection	3–4	3–4
Hand Tabulation/Discussion of Results	1–2	1–2
TOTAL	**8–10**	**7–9**

Determine Transportation Plan

Transportation should be provided to each surveyed area (cluster or lot), and possibly within surveyed areas. Drop-offs and pick-ups will need to be scheduled in advance. In urban or peri-urban areas, where families tend to reside more closely and within a smaller geographic area, transportation may not be an issue. The KPC coordinating team will decide whether special plans must be made to transport interviewer teams.

- If resources permit, it is recommended that each survey team have one available vehicle and a minimum of one vehicle per supervisor. Staff may share vehicles, take taxis or motorbikes, or use bicycles.
- Calculate the number of vehicles needed (to and within the project area). Teams in adjacent survey sites may share resources, vehicles, and supervisors.
- Map the survey sites to determine location and time/distance factors in getting to clusters. It is helpful to indicate important landmarks such as roads, schools, and churches, temples, or mosques on each map. Consider the map when determining the number and kinds of transport or vehicles needed.
- Calculate budget for fuel, maintenance, and drivers. Hire drivers who are familiar with the layout of the project area. Drivers can also serve other roles during the survey such as interviewers or guides, if necessary.
- Depending on travel factors, determine the minimum number of teams needed and the time required of supervisors and interviewers. If possible, arrange for the survey to be conducted in about 3 days.

When creating interviewer teams, remember that there are fewer supervisors than interviewers. To ensure high-quality data collection, each supervisor needs to be able to assist and monitor all interviewers on his or her team. A team of 3–5 persons (1 supervisor and 2–4 interviewers) is manageable.

Develop Plan for Editing, Printing, and Copying/Reproducing

- Plan to reproduce the following:
  - ⇨ Survey questionnaires (to be used during training and actual field implementation of the survey)
  - ⇨ Other training materials
  - ⇨ Additional field documents such as maps of survey sites, or a quick reference sheet summarizing household and respondent selection protocols for interviewers
  - ⇨ Tabulation tables
  - ⇨ Charts/graphs for presentations
  - ⇨ The survey report
  - ⇨ Other materials that will be developed to disseminate results to the community and relevant agencies/individuals.
- Determine available resources for editing, printing, and reproducing the survey questionnaire and other materials. This determination includes available power sources, computer hardware, software, qualified personnel, reproducing machines, and stores of paper, etc.
- Have a plan for developing and copying these necessary documents, including the following:
  - ⇨ WHO will develop and copy the documents
  - ⇨ WHERE can documents be developed and copied
  - ⇨ WHEN can documents be developed and copied.

Develop Plan for Tabulation/Analysis

- Decide how KPC questionnaires will be tabulated (manually and/or by computer), and plan personnel schedule and training accordingly
- Identify all relevant parties that can be involved in hand tabulating the data
- Assess the project's capacity for computerized data tabulation
- Determine resources—computers, printers, software, power sources, and availability of trained computer personnel

Develop Plan for Dissemination and Data Utilization

- When?
- Where?
- With whom?

- How?

Other Administrative and Logistical Issues

In some countries, projects have to receive special permission from the government to conduct the survey. Therefore, communicate with the proper officials early in the planning process to avoid future delays in the survey activities and/or negative responses to survey findings.

- Identify, reserve, and confirm facilities for training, tabulation, and feedback sessions
- Contact agencies/organizations with child survival, maternal health, or other relevant projects, and invite them to participate
- Schedule a briefing with the USAID mission and all other interested agencies (e.g., MOH) before the initial training activity
- Schedule a debriefing (feedback session) and a review of findings with all interested parties as soon as possible after the survey
- Obtain permission from appropriate administrative officials to conduct household interviews during the field test and the actual survey
- Purchase or borrow necessary supplies and equipment for facilitating training, conducting the survey, tabulation, analysis, and presentation of findings

3. DETERMINE NEEDS AND INFORMATION GAPS

One of the most important steps in preparing for a KPC survey is determining the information needs of the local context. Because of limited time and resources, projects are encouraged to use the KPC to gather information that:

- Cannot be found in existing data sources (or exists but is of poor quality)
- Can be used for either a) project planning and management or b) to estimate the effectiveness of project activities in achieving desired outcomes

CONSULT EXISTING PROJECT DOCUMENTS

Refer to existing documents such as the project's proposal or DIP to determine the following:

- Beneficiary population and related demographics
- Project goals and objectives
- Project indicators (in particular, those that can be measured in a survey)
- Which interventions to address in the survey
- Other relevant information on the target beneficiary population.

CONSULT WITH LOCAL PARTNERS AND STAKEHOLDERS

This activity was discussed in the previous section of the field guide. It is listed here to remind projects to engage local partners and stakeholders in the earliest stages of the KPC process.

CONSULT OTHER SOURCES OF DATA

In addition to determining the needs of the project and stakeholders, it is useful to identify what types of information already exist. Useful information may be found in the following:

- Demographic and Health Surveys (DHS) or other national surveys (e.g., Multiple Indicator Cluster Surveys (MICS))
- Health service statistics
- Other studies (e.g., in-depth surveys, qualitative research, data from other PVOs/NGOs operating in the same geographical area).

It is not uncommon for a PVO to discover that there are a number of studies conducted within

a country or a particular region of a country but little or no data on the specific project area. However, it still helps to refer to existing tools, methods, and results from other studies when designing your KPC survey. Once the KPC survey is completed, your project can compare findings from the KPC survey with existing estimates for the country or region as a whole.

Demographic and Health Surveys

DHS are nationally representative household surveys with large sample sizes. Many of the topics covered in DHS are also covered in the *KPC2000+*, which was modeled after the MEASURE DHS+ standard questionnaire. Some DHS surveys have used sampling designs that yield estimates for regions (or other sub-divisions) within the country in addition to national-level estimates.

PVOs are encouraged to consult DHS or other national surveys, which will be helpful when choosing what to include in a KPC survey. Once the data have been collected and analyzed, a project can also compare KPC findings with findings from DHS or other data to see how the project area relates to the entire nation (or a region, when such data are available) in terms of child health and survival.

FYI—For Your Information

Visit the MEASURE DHS+ Web site at **www.measuredhs.com**. The site includes a complete and up-to-date list of all DHS surveys. You can access DHS data immediately using the DHS *STATcompiler*. *STATcompiler* is an online database tool that allows users to select many countries and hundreds of indicators to create customized tables that serve their specific data needs. The *STATcompiler* accesses nearly all of the population and health indicators that are published in DHS final reports with just the click of the mouse.

Below is a list of recent DHS surveys. Visit the MEASURE *DHS* Web site for updated information.

RECENT AND UPCOMING MEASURE *DHS*+ SURVEYS (as of August 2001)		
Type of survey	*Country/Year*	*Dates of fieldwork*
DHS	Armenia 2000	Oct 2000–Nov 2000
SPA[1]	Bangladesh 1999	Jul 1999–Dec 1999
DHS	Bangladesh 2000	Oct 1999–Mar 2000
Special[2]	Bangladesh 2001	Jul 2000–Dec 2000
DHS	Benin 2001	Aug 2001–Oct 2001
DHS	Cambodia 2000	Feb 2000–Jun 2000
DHS	Colombia 2000	Mar 2000–Jul 2000
DHS	Egypt 2000	Mar 2000–May 2000
DHS	Ethiopia 2000	Feb 2000–Apr 2000
DHS	Gabon 2000	Oct 2000–Dec 2000
DHS	Haiti 2000	Mar 2000–Jul 2000
Benchmark	India 1999	-
DHS	Kazakhstan 1999	Jul 1999–Sep 1999
SPA[1]	Kenya 1999	Apr 1999–Aug 1999
DHS	Madagascar	Apr 2002–Jun 2002
DHS	Malawi 2000	Jul 2000–Nov 2000
DHS	Mali 2000	Jul 2000–Sep 2000
DHS	Mauritania 2000	Oct 2000–Dec 2000
SPA[1]	Mexico 2000	Jan 2000–Feb 2000
DHS	Namibia 2000	Sep 2000–Dec 2000
DHS	Nepal 2001	Jan 2001–Jun 2001
Special	Nicaragua 2001	-
Special	Niger 2000	Apr 2000–Aug 2000
DHS	Nigeria 1999	Mar 1999–May 1999
DHS	Peru 2000	Jun 2000–Sep 2000
DHS	Rwanda 2000	Jun 2000–Aug 2000
DHS	Senegal 1999	Oct 1999–Dec 1999
DHS	South Africa 1998	Feb 1998–Sep 1998
Interim	Tanzania 1999	Sep 1999–Nov 1999
DHS	Turkmenistan 2000	Jul 2000–Oct 2000
DHS	Uganda 2000	Sep 2000–Dec 2000
DHS	Zambia 2001	Aug 2001–Dec 2001
DHS	Zimbabwe 1999	Sep 1999–Dec 1999

[1]Service Provision Assessments (SPA) are surveys conducted in health facilities and communities to obtain information about the characteristics of health services such as their quality, infrastructure, utilization, and availability.
[2]Maternal Mortality and Maternal Health Services Survey.

Multiple Indicator Cluster Surveys

Since 1998, MICS have been conducted by UNICEF, in collaboration with a number of agencies and organizations, to assist countries in assessing their progress toward World Summit for Children goals. MICS assessments have been carried out in 100 less-developed countries (60 countries conducted stand-alone MICS studies and 40 countries incorporated some of the MICS modules into other studies). The MICS have provided valuable country-level data on global indicators related to child health and well-being.

FYI—For Your Information

You can access information on MICS by visiting **www.childinfo.org**. The Web site provides access to background information on the MICS, as well as the standard questionnaires, manuals, and a listing of countries where MICS studies have been conducted. Indicator data (by region and country) are also available on the Web site.

Health Service Statistics

Hospital or health center records, Community Health Worker registers, and other sources of health statistics often include information on topics such as the following:

- Immunization coverage
- Children's nutritional status
- The prevalence of childhood illnesses
- Prenatal, delivery, and postpartum care
- Child spacing
- Service use rates.

Projects should keep in mind that in areas where health service use is low, health service statistics may not reflect reality. In addition, health statistics do not provide information on processes or reasons why certain patterns exist. Nevertheless, such data offer a useful perspective when identifying problems.

It is possible that findings from a KPC survey, which is a community-based assessment, provide different results than health service data. In an effort to avoid major conflicts with the MOH, it is helpful to invite district/MOH staff to participate in the conducting and analysis of the survey. Projects are encouraged to be transparent in terms of how the data are collected and analyzed.

Other Studies

Before doing a KPC, explore what studies have already been conducted in the area or neighboring areas. For example, another PVO might have conducted a KPC survey in a neighboring district, or maybe a graduate student conducted a qualitative investigation

into local breastfeeding practices or delivery norms in select communities within the project area.

Because funding is often vertical, there is an inclination to only search for health studies. However, there may be studies in other sectors (e.g., education, water and sanitation, food security) that are relevant to your CS project.

As a reminder, always consider the specific information needs of your project. Even when there is some existing information for your project area, it might not relate specifically to the proposed activities of your project. Be creative in terms of modifying the generic KPC tools or using a sampling design that reflects the population groups targeted by your project. For example, it might be necessary to include grandmothers, males, and/or adolescents in your survey. Your tools and sampling strategy should reflect this design aspect. Consult with local experts (MOH statisticians, researchers from local universities or institutions) for assistance in designing a KPC questionnaire (or set of questionnaires, if there are different survey targets) and a methodology that best meets the needs of your project. Also consider the limitations of a KPC survey. Local experts can be helpful in identifying and designing ways to supplement your KPC data with information obtained from other methods (e.g., qualitative research techniques).

4. ADAPT AND TRANSLATE THE KPC QUESTIONNAIRE

Designing a survey tool that reflects the local context is critical to the success of a KPC survey. This section outlines the steps in developing an appropriate questionnaire.

CONDUCT FORMATIVE RESEARCH

Formative research is useful in designing a KPC survey. It allows a project to examine the social context of issues and can increase the validity and reliability of its surveys. Formative research can also identify local words and terms that are used to describe certain health phenomena (for example, AIDS, diarrhea, colostrum). Questions that contain context-specific information are more likely to be understood by respondents.

IMPORTANT ISSUES TO CONSIDER

You can use formative research to find out about the decisionmaking processes within households. Who makes decisions about maternal and child health and nutrition? Who determines treatment and/or care-seeking when a young child is ill?

Formative research can also be used to identify the individual(s) responsible for the care and well-being of young children. For example, in most cases, who is the primary caregiver of young children? What are the employment patterns in the target population, and what are the implications for child care (e.g., mothers might leave their young children with older siblings while they are away at work).

Answers to questions such as the ones listed above not only shed light on *what* to include in the survey, but also *who* to include.

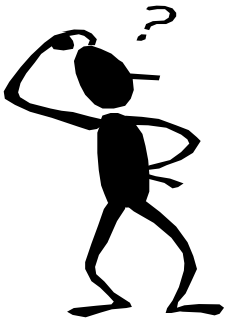
Qualitative studies use a wide range of methodological approaches. The specific techniques used in a particular study are chosen on the basis of the research question and setting but could include the following:

- ✓ Focus group discussions
- ✓ In-depth interviews
- ✓ Key informant interviews
- ✓ Observation of events or behaviors
- ✓ Ethnography
- ✓ Content analysis
- ✓ Role playing and story completion
- ✓ Systematic data collection (e.g., pile sorting, free listing, triadic comparisons, ranking).

Projects are encouraged to consider information-gathering activities, such as Participatory Rural Appraisal or Rapid Assessment Procedures, before conducting a KPC survey.

Consider using qualitative research after completing the survey. This method can help explain issues that were raised in the survey that could be more adequately explored in a structured interview.

ADAPT, TRANSLATE, AND BACK-TRANSLATE THE QUESTIONNAIRE



What should we include in our KPC survey?

The design and content of the questionnaire will depend on the needs of the project and stakeholders, which includes the types of respondents for the survey. For example, has the KPC coordinating team decided to interview mothers only, or is it also important to collect information on non-maternal caregivers? What about husbands/men of reproductive age? Other women of reproductive age? Adolescents?

A project can make minor changes to its sampling strategy (for example, by using parallel sampling) to collect information on different types of respondents. Sampling options will be described later in this guide. Keep in mind that interviewing different types of individuals requires a different set of questionnaires for each population sub-group.

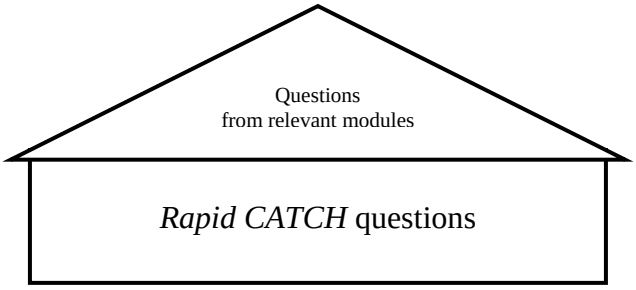
REMEMBER:

The decision about *who* to include in the survey will need to be made before deciding what to include.

Keep the following issues in mind when designing your survey:

- Clarify project interventions, objectives, indicators, and health messages.
  - ⇨ Are they current?
  - ⇨ Do they reflect national and international health education messages and standards?
- Use the Rapid CATCH questions as the basis for the questionnaire then add questions from the modules.

It is likely that some interventions covered by the *Rapid CATCH* questions are not addressed by a particular project. Consider the *CATCH* questions as the foundation of a house. *CATCH* questions reflect critical aspects of child health and survival that every project should consider. Use selected questions from the modules to distinguish your house from other houses in the CS community.



- ✓ Add questions that reflect the project's objectives, key indicators, and activities.
- ✓ Try to keep the questionnaire as short as possible.
- ✓ Keep skip patterns to a minimum.
- ✓ Do not ask "double-barreled" questions (asking about more than one thing in the same question).
- ✓ Phrase questions so that all individuals who are surveyed—highly educated individuals and individuals with little or no formal education—can understand the questions.

REMEMBER:

The questionnaire should include only those questions that the project will be able to use for making management or program decisions.

When deciding whether or not to include a question, the coordinating team should ask itself the following questions:

- What does the question mean, and why is it important? (Is this need-to-know versus nice-to-know information?)
- Is there a better way to collect this information?
- How will we use this information?

Translate and back-translate questions

Once the questions are selected, translate the questionnaire into the local language. When translating, pay attention to the following:

- ✓ Is the meaning of each question the same in the local translation as it is in the generic version of the questionnaire?

- ✓ Does the questionnaire reflect both the local language/dialect and the local context?

For instance, generic questions about foods should be translated to reflect locally available foods for relevant categories. If papaya is not a locally available fruit, it does not make sense to ask about it in the survey.

It is helpful to have one person or a group of people develop a local lexicon that lists local terms and phrases that can be used during interviews. It is particularly useful to list standard ways that questions can be re-phrased in the local language whenever a respondent does not understand the question that the interviewer is asking.

- ✓ Once someone has translated the questionnaire into the local language, a <u>different person</u> who is not familiar with the questionnaire should translate the questions back into the original language of the questionnaire (English, Spanish, French).

For example, if the local language is Quechua, then the questionnaire was probably created in Spanish and later translated into Quechua. Someone should translate the Quechua version back into Spanish to confirm that the translation was accurate.

Field-test (Pretest) Questions

- ✓ Before training interviewers, pretest the translated questionnaire by interviewing a small number of respondents (for example, 10 mothers). Field-testing can be done with the KPC field supervisors as part of their training and general orientation. This exercise not only familiarizes the supervisors with the questionnaire, but it provides an opportunity to further refine the questionnaire before it is duplicated and distributed to the interviewers. It also builds their capacity in critical thinking and questionnaire development.

- ✓ Assemble respondents and discuss their reactions to the survey. Find out:

  - What questions they did not understand
  - What questions seemed awkward or foolish
  - What their suggestions are to improve the wording of questions

- ✓ Modify the questionnaire based upon the field-testing experience. Keep in mind that additional changes might also arise during the interviewer training. Therefore, projects should have a clear plan in terms of how final revisions will be incorporated and how the final version of the questionnaire will be reproduced.

5. DESIGN A SAMPLING STRATEGY

Your PVO has recently been awarded a USAID CSHGP grant to implement a CS project in District X of a particular country. To identify problems, set objectives, and plan activities, staff members of the new project want to collect some baseline information on the target population. The Project Manager knows about cluster surveys but has just learned about other options at a recent workshop for health program managers. What factors should the KPC Coordinating Team consider when designing its sampling strategy? This section of the field guide reviews key sampling terms and concepts. It also presents guiding principles and decision trees to assist projects in designing a strategy that reflects the local context and the interests of project stakeholders.

The following guiding principles will be highlighted in this section of the field guide:

- Make random the standard
- Strive to be precise
- Include what you can use
- Use what you know
- The sky is not the limit.

Organization of this Section of the Field Guide

I. WHERE DO WE START?

When designing a survey, it is important to take aspects of the local context into account. If your PVO is new to the project area, you are probably unfamiliar with the local population. Local partners and stakeholders can help to identify information that should be considered when designing the sampling strategy as well as the questionnaire. The following are examples of characteristics to consider:

- Geographical distribution of the population
- Religious, language, and/or ethnic groups
- Castes/tribes
- Different household structures (for example, polygamous, female-headed, or child-headed households)
- Socioeconomic groups.

Child health and well-being might vary according to certain social, cultural, and economic characteristics, so it is important to consider whether some of those differences should be explored in the survey. Why is this important? Your project might want to document differences at baseline. Your project can use that information to better allocate resources and set different program targets for different segments of the population.

In addition to discussing aspects of the local context with stakeholders, look at existing data sources, such as censuses or community registers, to get a more complete picture of the target population.

It also helps to get a visual image of the program area. Find out if there are maps of the local area. If not, consider preparing one. Map where communities are located and, if possible, note major subdivisions (such as wards or health facility catchment areas). It also helps to include major landmarks (roads, rivers, health facilities, places of worship, schools, and markets) on the map.

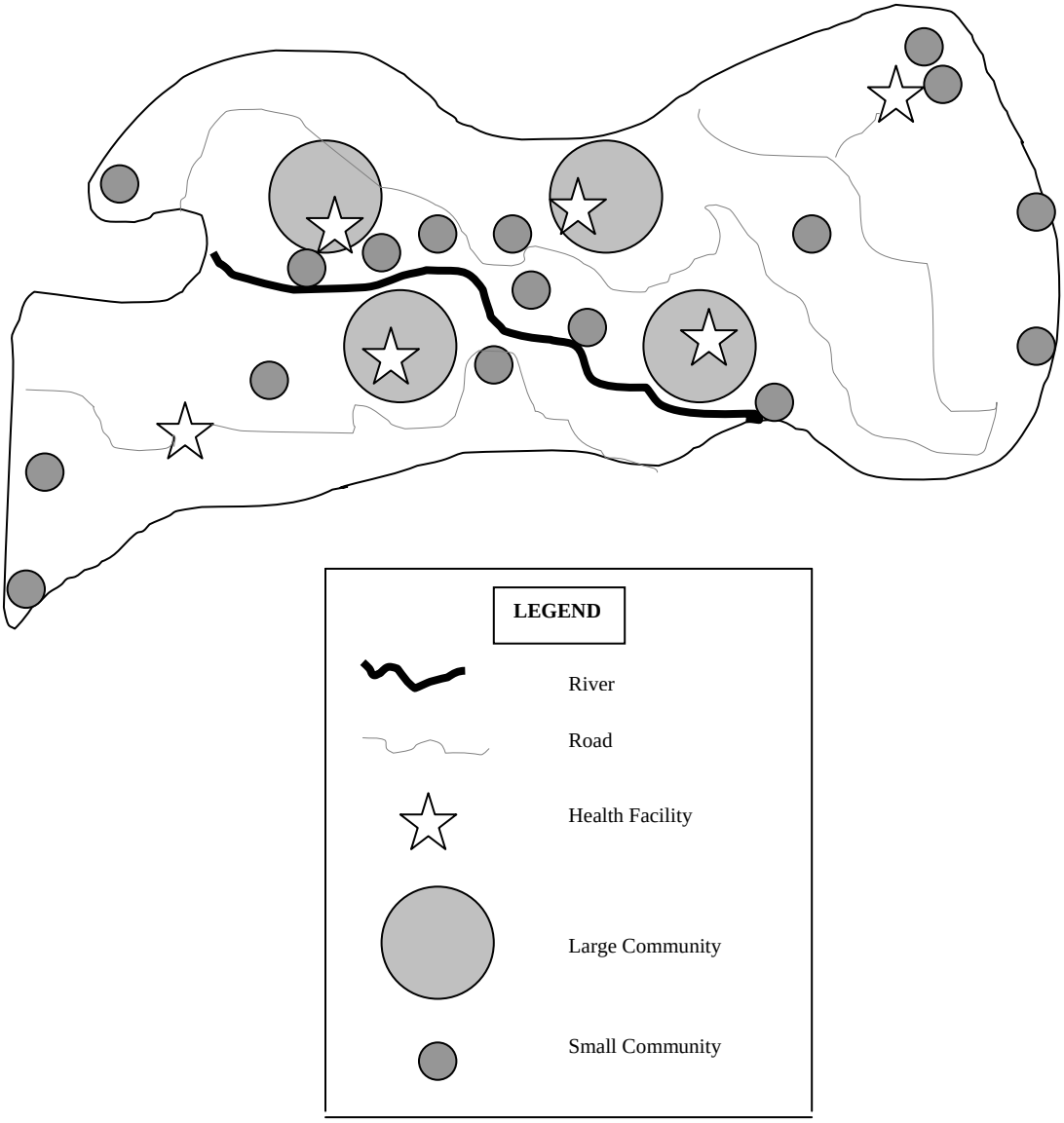
The following page shows a map of a fictitious district (District X). To keep the illustration simple, the map does not include landmarks other than roads. However, by examining the map, one can get a visual sense of how the population is distributed throughout the district.

As seen on the map, most of the communities are located near the river. The district has four large communities and some remote villages at the edges of the district. At a KPC planning meeting, local stakeholders express a desire to collect information on remote villages. In addition, someone mentions that the HIV/AIDS epidemic has taken a toll on the community, which now has many children being raised by persons who are not their biological parents. Should the project hire enumerators to visit

every household so that information is collected on important segments of the population?

That would be very costly and time consuming, and the project has a limited amount of time and resources to collect its baseline data.

Illustrative Map of District X



How can the project balance local information needs with the fact that it has limited resources? The answer is *sampling*. (See Table 5.1 for definitions of key sampling terms.) A project can collect information from a *sample* (a subset of the population) rather than from every person. Sampling units can be individuals, households, or communities, depending on the focus of the study.

Table 5.1
A Review of Key Terms and Concepts

TERM/CONCEPT	DEFINITION
Bias:	An error that consistently results in an over or under estimation of a value of a measurement. Bias can result from problems with how the sample was selected. Use of a random and/or systematic sampling process may help prevent this "selection bias."
Cluster:	A naturally occurring group of individuals that is likely to include a specified number of individuals from a population group of interest.
Cluster sampling:	A method of sampling population clusters rather than individuals, then interviewing a certain number of individuals within each cluster to achieve the desired sample size.
Confidence interval (limits):	Indicates the range of possible values that the sample estimate will fall within a certain percentage of the time. *Confidence limits* are the highest and lowest values within that range, and are usually calculated at a level of 95 percent. That is, there is a 95 percent chance that the actual rate or proportion being estimated in the study falls within the confidence interval.
Cumulative:	Increasing a sum by continuing to add to it. For example, assume there is a list of three communities. Community A has 40,000 people, Community B has 60,000 people, and Community C has 50,000 people. The cumulative population of Community A and Community B is 100,000 (40,000+60,000). The cumulative population of Community A, Community B, and Community C is 150,000 (40,000+60,000+50,000).
Lot Quality Assurance Sampling (LQAS):	A special form of stratified sampling that allows projects to identify areas with levels of coverage that are at or above expectation versus those that are below expectation.
Multi-stage sampling:	A process involving more than one step of sampling before reaching the ultimate unit of interest. For example, with cluster sampling, projects first sample clusters from the population, then households within clusters, and finally, mothers/caregivers within sample households.
Probability proportional to size (PPS):	A sampling principle that ensures that the sample's distribution mirrors the population's distribution. Communities with larger populations have a proportionately greater chance of having clusters located in those communities than communities with smaller populations.
Random sample:	A method of selecting a sample that ensures that each unit in the population has an equal chance of being selected.
Random number:	A number that is selected (by chance) from many numbers. Each number has an equal chance of being selected.
Sample:	A group of units (such as individuals or households) selected from the general population.
Sample area:	Community (cluster, lot) selected from the general population for a study.
Sample size:	Number of units (individuals, households) selected from the population for inclusion in a study.
Sampling unit:	Usually the same as the unit of analysis. It is the unit from which information is collected in a survey. For KPC surveys, the sampling unit is usually the individual or the household. However, these units can sometimes be aggregated to reflect community-level phenomena.
Sampling frame:	List of every possible sampling unit within the target population from which a sample will be drawn.
Sampling interval:	The total population size (N) divided by the sample size (n). Used as part of systematic sampling to select units from a sampling frame
Standard error:	Also known as *sampling error.* It is a statistical measure that indicates the precision of a sample estimate, and is used to calculate the confidence limits of that estimate.
Systematic sampling:	A sampling approach that involves calculating a *sampling interval* based on the required sample size. A random starting point is chosen, and then cases are selected from the sampling frame at a sampling interval.

Sometimes there is a need to sample different units at different stages of the sampling process to get to the ultimate units of interest. In the case of the KPC, project beneficiaries (children less than 24 months and their caregivers) are the ultimate units of interest. However, to get to these units, a project must

sample at the following stages:

STAGE 1: Sample communities from the population to get sample areas (clusters)

STAGE 2: Sample households within sample areas

STAGE 3: Sample mothers/primary caregivers of children less than 24 months in the sample households.

Collecting information from a sample is much more efficient (in terms of the amount of time and money spent) than collecting information from every unit in the general population. If the sample is selected *randomly,* findings from the sample should reflect what is going on in the larger population. However, if the project wants to collect information from certain population groups, it can modify its sampling design to make sure that individuals from those groups are included in the study.

Guiding Principle #1: Make random the standard

II. SIMPLE RANDOM SAMPLING

When you randomly select units from the general population, you ensure that every unit has an equal chance of being included in the study. Random sampling involves selecting units based upon chance. (See Table 5.2 for examples of how to select units based on chance.) If you do not select a random sample, your results might be biased.

In theory, *simple random sampling (SRS)* is the ideal way to select units from the general population. However, most studies do not use SRS. One reason is because SRS requires a *sampling frame* (a listing of every unit in the population). With SRS, every unit in the sampling frame is assigned a unique number, and then, a sample is drawn by randomly selecting numbers until you reach the desired sample size. For example, if your project wants to randomly select 300 households, it will need to 1) list every household in the project area, 2) assign a number to each household, and 3) randomly select 300 numbers. If the project wants to select 300 individuals, it will need to 1) list every person who lives in the project area, 2) assign a number to each individual, and then 3) randomly select 300 numbers (corresponding to 300 different people).

What makes a good sampling frame? Any complete and up-to-date listing of all units in the total population can be used. The following are some examples:

- Census (either pre-existing or conducted by the PVO)
- Voter registration lists
- Tax lists
- Community health worker registers

- Surveillance records
- Maps of the area showing each dwelling.

In most cases, a good sampling frame will not be available. If a sampling frame is not complete, every unit in the population does not have an equal chance of being selected for the study. As a result, using a poor-quality sampling frame might introduce a bias (a systematic error) into the study. The presence of a bias affects whether sample estimates accurately reflect rates in the general population.

Table 5.2

HOW DO WE SELECT RANDOMLY?

Without random sampling, there is a risk that the sample you selected is not representative of the broader population. Below are some ways to select units randomly. The KPC Coordinating Team should decide on which method(s) to use and instruct all survey teams to use the same method(s).

WITH A SAMPLING FRAME

When you have a complete and up-to-date sampling frame, you can use one of the following to select units randomly. Remember to number each unit in your sampling frame.

A. **Currency note**—Get a local currency note and look at the last few digits of the serial number. Usually, there are more digits in the currency note's serial number than in your sampling interval. To select a random number, refer to the number of digits in your sampling interval. This will determine how many digits you will refer to in the serial number. For example, if you have a two-digit sampling interval, look at the last two digits of the serial number; with a three-digit sampling interval, look at the last three digits of the serial number, and so on. As an illustration, suppose your sampling interval equals 750. You have a currency note with the serial number 123456789. Since you have a three-digit sampling interval (750), look at the last three digits of the serial number (in this example, 123456**789**). If the last three digits are less than or equal to your sampling interval, use that number as your random number. If the last three digits of the serial number are greater than the sampling interval (in the above example, 789 is greater than 750), select a new three-digit number by shifting to the left by one digit. In this example, that new number is 678 (12345**678**9). If the new number is less than your sampling interval (in this case, 678 is less than 750), use that number as your random number. If no set of three digits in the currency note is less than or equal to your three-digit sampling interval, then you will need to use another currency note to identify another number. Please note, the random number can include zero (for example, 079).

B. **Random number table**—Some statistics books include a table of random numbers. To use a random number table, get a pen or pencil, close your eyes, then let the point of the pen/pencil land on the table. The number that the pen/pencil lands on will be your random number. If the random number has more digits than the sampling interval, you can use an approach similar to the process described for currency notes.

C. **Random number generator**—Some computers and calculators have random number generators that will choose random numbers for you.

Table 5.2 (continued)

HOW DO WE SELECT RANDOMLY?

WITHOUT A SAMPLING FRAME

When you do not have a sampling frame, you can use one of the following options:

- Flip a coin
- Select using slips of paper
- Spin the bottle
- Divide the area into smaller and smaller sub-areas.

Options A and B are helpful when you have a small number of choices, whereas options C and D are appropriate when you have a large number of choices. Options C and D are particularly useful when selecting a starting household within a sample area (that is, a cluster or lot).

A. **Flip a coin**—When a decision only involves two choices (such as "Go left or go right?" or "Visit house A or house B?"), you can flip a coin to make sure that the selection process remains random. Be sure to decide which side represents which option before flipping the coin.

B. **Use slips of paper**—When there are only a few options, you can write those choices on small slips of paper, and then randomly select one slip of paper. For example, if you are trying to decide which of 10 houses to use as your starting point, number the houses from 1–10 and write each number on a small slip of paper. Place all of the slips in an envelope, sack, or other container, and then remove one of the slips. The house that corresponds to the number that you selected will be your starting point.

C. **Spin the bottle**—The spin the bottle technique has been used widely in KPC surveys to identify the starting point within a sample area. Spinning a bottle (or a ballpoint pen) at the center of the community helps the survey team randomly choose a direction to follow. To select by spinning a bottle, do the following:

    1. Go to the population center of the sample area (the point in the community where the population is about equally distributed on all sides).
    2. Select a smooth, level spot where you can place the bottle.
    3. Spin the bottle.
    4. When the bottle stops spinning, determine which direction the mouth of the bottle is pointing. The survey team should go in the direction that the bottle is pointing.

*NOTE: If two survey teams are conducting interviews in the same area, they should go in opposite directions. In other words, one team goes in the direction selected by the spinning bottle, and the other team should go in the opposite direction.

Table 5.2 (continued)

HOW DO WE SELECT RANDOMLY?

WITHOUT A SAMPLING FRAME (continued)

D. **Divide the Sample Area into Smaller and Smaller Sub-Areas (Selecting by Subdivision)**—Although the spin the bottle technique is very popular, selecting through subdivision is a more rigorous way of randomly selecting the first household in a sample area. This method of subdivision is helpful if a team wants to choose a subsection of a very large community to conduct interviews. It might take longer than the spin the bottle technique, but it is a more desirable method, particularly if the population in the sample area is not spread out over a large area. Below are the steps to select by subdivision.

1. Go to the population center of the sample area and identify four directions (north, south, east, and west) or four sections (quadrants). There is usually a landmark—a church/mosque/temple, market, or school—at the center of the community. If it is not clear where the center of the community is, look at a map of the area. Also, meet with persons who are familiar with the sample area to help you identify the center of the community.

2. Write the four sections/directions on four slips of paper and put them in a container. Randomly select one of the slips of paper.

3. Proceed in the selected direction and go to the place that equally divides the quadrant's population in half. In other words, find out where 50 percent of the quadrant's population resides on one side and the other 50 percent resides on the other side. Randomly select (for example, by flipping a coin) which of the two ways to proceed. *Repeat this step until you have a small and manageable set of houses, then proceed to Step Four.*

4. Count all the households in the area that resulted from the process described in Step Three.

5. Use a random number table to select the first household where an interview will be conducted.

Determining whether or not there is a good sampling frame is only the first step. Your project will also need to decide how many units from the general population to include in the study. In other words, you will need to calculate the *sample size*.

Table 5.3 presents the sample size formula for simple random samples.

TABLE 5.3

Calculating the Size of a Simple Random Sample

The sample size formula for a simple random sample is as follows:

$$n = z^2 (pq)/d^2$$

where **n**= sample size
z= statistical certainty chosen
p= estimated level/coverage to be investigated
q= 1-p
d= precision desired

Usually, the value of **p** is not known, in which case you can be conservative and choose p=0.5. This value will give you the largest possible sample size for your study.

The value of **d** depends on your desired level of precision and should be chosen according to the objectives and needs of the survey. A precision of 10 percent (d=0.1) is widely used and is acceptable if your project seeks information for project management purposes.

When assigning the value of **z**, most studies assume a 95 percent confidence level. A 95 percent confidence level means that there is a 95 percent chance that the true rate in the population is within the range of values defined by the confidence limits of your survey's estimate. The corresponding **z** value for a 95 percent confidence level is 1.96.

Using the above values, the sample size needed for a random sample survey is as follows:

$$n = (1.96)^2 (.5 \times .5)/(.1)^2$$
$$n = (3.84)(.25)/(.01)$$
$$n = 96$$

A NOTE ABOUT SAMPLE SIZE

The formula presented in Table 5.3 calculates the sample size for one survey at one point in time. Projects that are interested in comparing changes over time (for example, by comparing baseline and final surveys) will need to increase the sample size of each survey. This rule applies regardless of the type of sampling design used and will be discussed in more detail later on in the field guide.

III. ALTERNATIVES TO SIMPLE RANDOM SAMPLING

Cluster Sampling

When deciding if your project should use SRS, sample size and resource (time, money, personnel, etc.) requirements are just as important as whether or not there is a sampling frame. To achieve a sample size of 96 (as calculated in Table 5.3), a project will have to repeat the random selection process 96 different times. In addition, those 96 units (households/individuals) might live great distances apart, making it difficult for survey teams to complete all of the interviews within a short time

frame. Are there more efficient ways to select a sample? One possible alternative is *cluster sampling*. With cluster sampling, clusters (not individuals) are randomly selected from the general population. Several individuals within each cluster are then interviewed to reach the desired sample size.

WHAT IS A CLUSTER?

A *cluster* is a naturally occurring group of individuals (such as a village, ward, or city block) likely to include the population group your project is interested in studying. In the case of the KPC, your project is probably interested in children less than 24 months and their caregivers.

Cluster sampling has been used worldwide to assess coverage of the World Health Organization's (WHO) EPI. Within the CS field, many people think that the terms "KPC" and "cluster sampling" mean the same thing. This is due to the fact that most CS projects have used cluster sampling with their KPC surveys.

How did cluster sampling become such a popular method? Below are two reasons.

1. It does not require a sampling frame, other than a list of population centers (such as towns, villages, or communities) and their estimated population sizes.

2. By interviewing a number of people who live in the same cluster, it reduces time and travel costs between interviews.

In sum, cluster sampling is an efficient way to get coverage estimates for an entire program area.

With cluster sampling, clusters are usually selected using *systematic sampling*. To use *systematic sampling*, a project needs 1) a sampling frame of all communities in the program area and 2) a sampling interval. The sampling frame should include every community in the project area, its population size, and its cumulative population (see Table 5.6 for an example).

Cluster sampling, like SRS, is a *probability sampling* design. In other words, the selection of sampling units is based on chance. Therefore, it is important to list every community in the program area, regardless of its size. Clusters are selected with *probability proportional to size* (PPS). This means that larger communities have a greater chance of having clusters than smaller communities. Why sample with PPS? This is done if you want the distribution of cases (mothers/caregivers and young children) in the sample to mirror the distribution of cases in the general population. Table 5.4 outlines the steps for sampling with PPS.

Table 5.4

STEPS FOR SAMPLING CLUSTERS WITH PPS

1. Calculate the sample size based on the desired level of precision and confidence. Most KPC surveys use a sample size equal to 300.

2. Determine the number of interviews per cluster. It is suggested that your project aim to conduct 10 interviews in each cluster.

3. Divide the sample size by the number of interviews in each cluster. This will give you the number of clusters. If you plan for a sample size of 300 and a cluster size of 10 interviews in each cluster, you will have 30 clusters in your survey.

4. Refer to existing population data to get the size (number of residents) of each village/town/ward in the program area.

5. Calculate the cumulative population of each village/town/ward by summing the total population of the village with the combined total population of all the preceding villages on the list (see Table 5.6). NOTE: the cumulative population of the last community listed in the sampling frame should equal the total population of the entire program area. If this is not the case, check your calculations.

6. Determine the sampling interval by dividing the total population of the entire program area by the total number of clusters.

7. Choose a random number. This number will be used to identify the starting point on the list to begin selecting clusters. The random number must be less than or equal to the sampling interval. As an example, if the sampling interval is 10,039 (see example in Table 5.5), you would select a random number between 1 and 10,039. As an example, assume that you used one of the techniques presented in Table 5.2 for selecting a random number, and chose 9,679 as the random number.

8. Look at the column where you have listed the cumulative population of each community and determine which community contains (that is, the cumulative population equals or exceeds) the random number. This is Cluster #1. In Table 5.6, Utaral (the first community listed in the sampling frame) has a cumulative population that equals or exceeds the random number chosen in STEP 7.

9. To identify the second community where a cluster is located, add the sampling interval (10,039) to the random number selected in STEP 7 (9,679). The community whose cumulative population equals or exceeds that number is the location of Cluster #2. Using the data in Tables 5.5 and 5.6, Cluster #2 is located in Talum because 10,039+9,679=19,718, and the cumulative population in Talum includes that number.

10. To identify the remaining clusters, add the sampling interval to the number that identified the location of the previous cluster.

*NOTE: A community can contain more than one cluster.

Table 5.5
Calculating a Sampling Interval

Formula for calculating a sampling interval:		
SAMPLING INTERVAL =	Total population to be surveyed / Number of clusters	
*****EXAMPLE (based on data in Table 5.6)*****		
A=	TOTAL POPULATION IN THE PROGRAM AREA	= 301,170
B=	TOTAL NUMBER OF CLUSTERS IN THE SURVEY	= 30
C=	A÷B (301,170/30)	= 10,039
*It is okay to round the sampling interval to the nearest whole number. For example, if the sampling interval calculated above was equal to 10,039.3 you would round to 10,039. If it was equal to 10,039.5, round up to 10,040.		

Table 5.6
Using Systematic Sampling to Select 30 Clusters with PPS
*Data taken from *Training for Mid Level Managers: The EPI Coverage Survey (WHO, 1991)*

ASSUME RANDOM NUMBER=9,679; SAMPLING INTERVAL=10,039

No.	Name of Community	Population	Cumulative Population	Cluster	No.	Name of Community	Population	Cumulative Population	Cluster
1.	Utaral	12,888	12,888	1	26.	Nozop	17,808	157,117	14,15
2.	Bolama	3,489	16,377		27.	Mapasko	3,914	161,031	16
3.	Talum	6,826	23,203	2	28.	Lothoah	14,006	176,037	17
4.	Wara-Yali	4,339	27,542		29.	Voattigan	9,584	185,621	18
5.	Galey	2,203	29,745		30.	Pliotok	4,225	198,846	19
6.	Tarum	4,341	34,086	3	31.	Dopoltan	2,643	201,489	20
7.	Hamtato	1,544	35,630		32.	Coccopa	26,000	227,289	21,22
8.	Nayjaff	885	36,515		33.	Famezgi	3,963	231,452	23
9.	Nuviya	2,962	39,477		34.	Jigpelay	2,115	233,567	
10.	Cattical	4,234	43,711	4	35.	Mewoah	507	234,074	
11.	Paralal	1,520	45,231		36.	Odigala	3,516	237,590	
12.	Egala-Kuru	3,767	48,998		37.	Sanbati	14,402	251,992	24,25
13.	Uwanarpol	3,053	52,051	5	38.	Andidwa	2,575	254,567	
14.	Hilandia	60,000	112,051	6,7,8,9,10,11	39.	Ore-Mikam	3,105	257,672	
15.	Puratna	2,207	114,348		40.	Dunu-Mikam	4,176	261,848	26
16.	Kagaini	1,355	115,703		41.	Kedi-Sina	1,919	263,767	
17.	Hamali-Ura	833	116,536		42.	Panabalok	3,261	267,028	
18.	Kameni	4,118	120,654	12	43.	Rokini	4,270	271,298	27
19.	Kiroya	2,782	123,456		44.	Talosso	3,301	274,599	
20.	Yanwela	3,285	126,721		45.	Djaragna	3,250	277,849	
21.	Bagvi	4,416	131,137	13	46.	Bibachi	4,670	282,519	28
22.	Atota	3,188	134,325		47.	Bilam	757	283,276	
23.	Kogouva	1,179	135,504		48.	Sisse	12,037	295,313	29
24.	Ahekpa	612	136,116		49.	Anda-Dali	2,155	297,468	
25.	Yandot	3,193	139,309		50.	Varok	3,702	301,170	30

Going back to the illustrative map at the beginning of this chapter, there are four large villages in District X. Suppose that 60 percent of the population in District X lives in those four communities. If we sample with probability proportional to size, we will select the majority of clusters from those four villages, which will result in a sample distribution that is similar to the population distribution in the district. If there is a particular interest in collecting information from small, remote villages (for example, if stakeholders claim that children in those villages are more at risk for poor health outcomes), then the project can modify its sampling strategy (for example, by dividing the project area into two groups: small communities and large communities, and then selecting a certain proportion of your respondents from small communities) to make sure that there are enough cases from remote villages.

IMPORTANT ACTIVITIES AFTER CLUSTERS ARE SELECTED

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

Community leaders can also provide useful information in terms of the layout of households within the sample area. It helps to draw maps of each sample area with the locations of each household. Survey teams can use this to at least select the starting household. If it is possible to identify all households with children less than 24 months, then the village leader can work with the survey team to randomly select 10 households to visit. Ask the village leader to encourage the participation of selected households in the survey. Your local partners, who might already have a rapport with the target communities, can be very helpful in communicating with the target areas, particularly if your project is new to the geographic area.

In addition to issues related to cluster selection, other sampling issues will arise when selecting households. The process of establishing the correct starting point in each cluster is critical because a cluster is formed from the first household selected. Therefore, any bias resulting from how that household was selected might invalidate the entire cluster. It is important to avoid selecting a starting household because of factors like convenience or the identity of the owner. Remember: random selection is the standard.

How can your project stay true to the random selection process? Randomly sampling from a large area can be both expensive and difficult. However, if it is possible to subdivide a large area into areas of smaller size (geographically and/or in terms of the number of residents) it becomes more practical to select a random sample. Refer back to Table 5.2 for information on how to select by subdivision.

Consider natural subunits within the cluster such as kindreds, paras, wards, or other subdivision units. There might not be a sampling frame for the project area as a whole, but in most instances there is some way of keeping track of households in smaller divisions within the project area. Within a particular community, the village or tribe head probably has a way of keeping track of members of the village/tribe. The supervisor can meet with the village head to find out if a list exists, or they can map out the general location of all households. To encourage ownership and consensus, the supervisor could ask the village leader to randomly select households with children under 2 by drawing slips of paper from a bag. The leader could also assist the survey team in arranging the interviews. **Remember to keep the process of selecting households random**. Interviewers should not visit a particular household just because a community leader says they should.

SUGGESTION:

The selection of the first household is a critical and sometimes time-consuming step in the KPC process. Consider having survey teams visit each sample area a day before starting interviews to correctly identify the starting household within each sample area.

After identifying the starting household in a cluster, most projects use the "nearest door" rule when selecting the remaining nine households in that cluster. In other words, after completing an interview, interviewers visit the household whose front door is closest to the door of the household they just visited.

The decision trees at the end of this chapter will guide your project in developing protocols for selecting multiple clusters (sample areas) within the same community, as well as in selecting households and respondents within each sample area.

WHAT IS A HOUSEHOLD?

A household is a group of people sharing the same kitchen, cooking area, or cooking pot.

There can be many nuclear units (each consisting of a man, his wife, and their children) within the same household, but if these units share the same cooking area, it is likely that they also share the same knowledge and behavior. If information is collected on more than one child in the same household, you will probably over represent practices and behaviors (for example, infant feeding practices, care-seeking practices) of that particular household in your sample. As a result, interviewers should collect information on only one child within each selected household.

Shortcomings of Cluster Sampling

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

- Its inability to provide information on subdivisions within the program area
- The design effect

A cluster survey provides a project with coverage estimates for the entire program area. If your project wants coverage estimates for different management units (for example, supervision areas or health-facility catchment areas) in the program area, it will need to explore other sampling strategies.

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

The sample size formula in Table 5.3 resulted in a sample size of 96. Multiplying 96 by 2 equals 192. For EPI 30-cluster surveys, the sample size is increased to 210 because it can be easily divided between 30 clusters (7 interviews in each cluster). For KPC surveys, the sample size is further increased to 300 (10 interviews in each cluster) because KPC surveys are used to estimate coverage for many different technical interventions, not just for immunization. A sample size of at least 300 is usually adequate for looking at sub-samples (such as children 0–5 months to assess exclusive breastfeeding), because the objective is to use KPC information to make management and programming decisions.

Improving on the Traditional 30-Cluster Design

As mentioned before, the traditional 30-cluster sampling design is widely accepted and is an efficient way to collect program-wide data. However, if your project wants to generate more precise and/or programmatically meaningful data, it can make changes to the cluster design, or choose another methodology altogether.

Guiding Principle #2: Strive to be precise

Surveys give us estimates of actual rates or proportions that exist within the general population. Because a project is estimating rates based on a sample, it should never regard an estimate from the KPC as an absolute number. Each sample estimate falls within a range of possible values. This range is defined by the *confidence limits*. Confidence limits indicate the margin of error associated with a sample estimate. When we estimate a rate or proportion in a study, we expect the confidence interval to include the actual rate *most of the time*. How often is most of the time? Phrases like “95 percent confidence” give us an idea. When you calculate 95 percent confidence limits, you assume that there is a 95 percent chance that the actual rate

being estimated by the survey falls within the confidence interval. The terms "confidence limits" and "confidence interval" are often used interchangeably. A confidence interval is the range of possible values, whereas the confidence limits are the lowest and highest values within that range.

The following formula can be used to calculate the confidence limits of each estimate:

$$\mathbf{P = p \pm z_a \sqrt{(pq)/n}}$$

where: P = the actual rate/proportion in the general population

p = the survey estimate

q = 1-p

z_a = the confidence level (with a 95 percent confidence level, this value equals 1.96)

n = sample size

The value for ($\mathbf{z_a \sqrt{(pq)/n}}$) is the margin of error. As an example, assume that your KPC survey gave a point estimate of 40 percent for the percentage of children with diarrhea who received ORT, and that the margin of error for that estimate was $\pm$ 9 percent. This means that although 40 percent is our best estimate, the actual rate of ORT use may be as low as 31 percent (40- 9) or as high as 49 percent (40+9).

Although projects should expect some variation in survey estimates, they should aim to be as precise as possible. Confidence limits are determined by a statistical measure known as the *standard error* (also known as *sampling error*). A precise estimate is one that has a small standard error (and therefore a narrow confidence interval). A less-precise estimate has a larger standard error and a wider confidence interval. These are important concepts to understand because if precision is high, there is a greater chance that estimates from the sample reflect what is really going on in the population as a whole.

Precision generally increases as the sample size increases. However, it is also important to remember that precision can vary by indicator. That is because some KPC indicators are limited to a smaller number of cases (subsamples), whereas others are based on a larger number of cases.

The KPC was not originally designed to measure improvements in the target population over time, but rather to do the following:

- Identify local problems at baseline so that the project can set program objectives
- Assess whether program objectives have been met (usually done at the end of project).

A sample size calculation of 300 takes into account the lack of precision of one estimate in time. Some projects might want to increase the precision of KPC estimates to measure changes over time. If a project will need to compare baseline results with final results, it will have to consider the lack of precision in estimates from two different points in time. When comparing two sets of estimates, you do not want their confidence intervals to overlap. If the intervals overlap, your project cannot be certain that a significant improvement took place over time. Therefore, to measure changes over time, your project would need to either increase the sample size of each KPC survey or bring about large changes in coverage through very focused interventions.

SAMPLE SIZE CONSIDERATIONS WHEN MEASURING IMPACT

To demonstrate impact, it is important to rule out factors that are not related to the project but that could have an effect on population outcomes. To determine the impact of their activities, some projects collect data on control or comparison groups (communities that are not beneficiaries of the project activities but who are otherwise similar to the communities that are being targeted by the project).

There are many types of control groups. Regardless of the type of control used, a project will have to compare two sets of estimates: one from communities exposed to the program's interventions and the other from communities who were not exposed. The same issues that arise when comparing estimates from two points in time should also be considered when a project wants to use control group evaluation designs.

**For more information on evaluation designs, refer to the article by Habicht et al. (1999) listed under References.*

As mentioned before, precision increases with increasing sampling size. You can increase the sample size of a cluster survey in one of two ways:

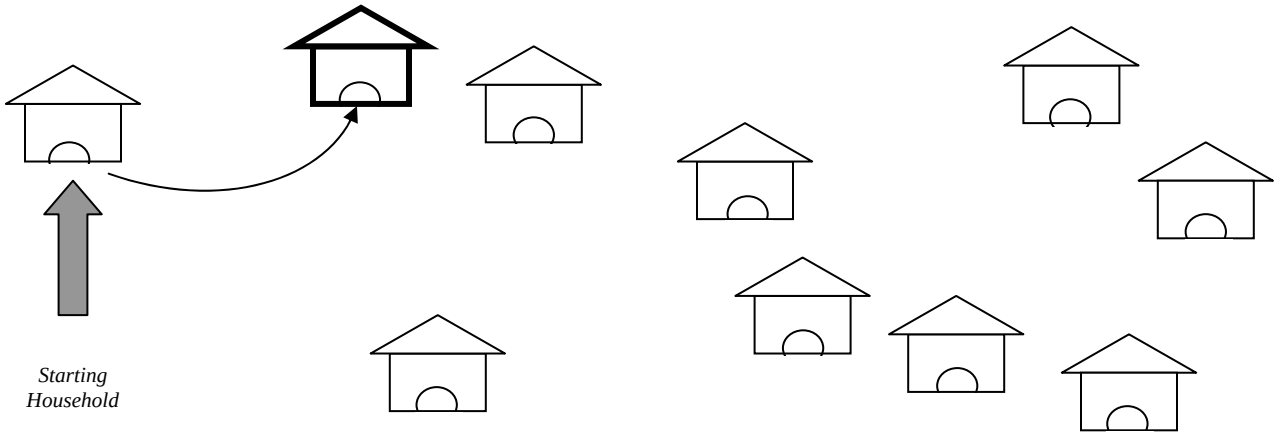
- ✓ Increase the number of clusters
- **x** Increase the number of interviews conducted in each cluster

Increasing the number of clusters is much more desirable than increasing the number of interviews in each cluster. That is due to one factor: the design effect. As mentioned before, individuals living in the same cluster are likely to share the same behaviors, practices, and other traits. As you increase the number of interviews in each cluster, you are likely to include more people who are alike in terms of the knowledge and practices your project is interested in. Therefore, increasing the number of interviews in each cluster increases the design effect. This results in survey estimates that are less precise and accurate.

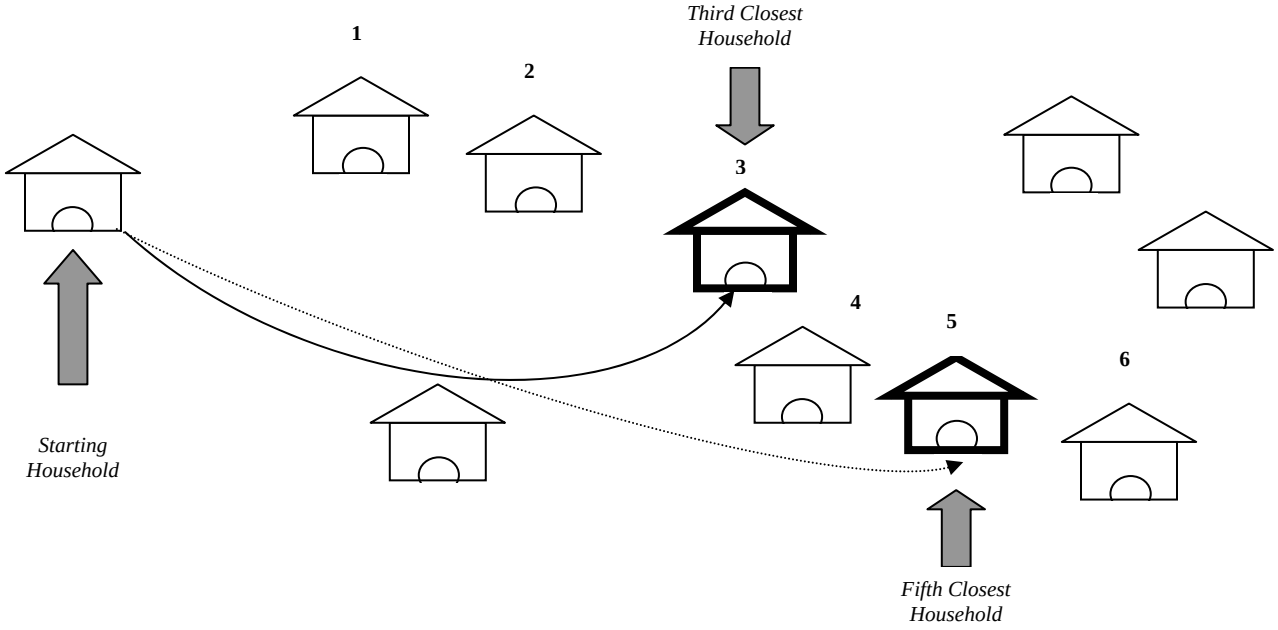
By reducing the homogeneity within each cluster, a project can increase the precision of its survey estimates. One way to do this is to change the protocol for sampling households. For example, instead of instructing interviewers to visit the nearest household, your project could instruct interviewers to visit the third closest or fifth closest household (see illustration below). If you sample households farther apart,

they might be less likely to share the same knowledge and practices. Therefore, by using the “third closest” or “fifth closest” rule, your interviewers are visiting a broader area of the community. This could help to reduce the design effect.

VISITING THE NEAREST HOUSEHOLD



VISITING EVERY THIRD OR FIFTH HOUSEHOLD



There is another way to reduce bias that may result from the selection of households within a cluster. If each survey team is able to meet with the village or tribal leader and create a sampling frame (if one does not already exist), then interviewers can randomly select 10 households from that sampling frame.

REDUCING HOMOGENEITY AT WHAT COST?

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

Even though it is better to increase the number of clusters, this will also increase survey costs. Unless there is a reason to select more than 30 clusters, it is probably not worth the effort. Why would a project go through the extra trouble of collecting data from additional clusters? One major reason is to be able to make comparisons between different parts of the target population. For example:

- To compare rural and urban/peri-urban communities
- To compare different districts within the same program area
- For projects that are expanding the reach of program activities to additional communities: to compare newly added communities with communities that are already being target by the project

In the scenario presented at the beginning of this chapter, local stakeholders of the fictitious project in District X expressed a desire to include remote villages in the baseline assessment. The project could respond to this request by conducting a certain proportion of KPC interviews in remote areas, and then comparing these data with data from cases in more centralized communities.

In instances such as the ones listed above, it makes sense to have an equal number of clusters in each stratum (group). Can a project choose to maintain the 30-cluster design, selecting 15 clusters from one group and 15 clusters from the other? Although 150 cases (10 interviews x 15 clusters) in each group seems like a fairly large number, it is important to remember that many KPC measures are based on subsamples. For example, in a sample of 150 children 0–23 months, there might only be between 30 and 40 children 0–5 months. Because you do not want to compromise the precision of certain estimates, it is best to increase the total sample size of the survey beyond 300 cases.

It is not always necessary to select 30 clusters from each group that you are interested in comparing. Your project will need to decide the minimum difference that

it would like to discover between groups, and then calculate the sample size in each group using the following formula:

$$n = z^2 (p_1q_1 + p_2q_2)/d^2$$

where: n = the sample size in each group
z = statistical certainty (for a 95 percent confidence level, z=1.96)
p_1 = the estimated proportion in group one
q_1 = 1- p_1
p_2 = the estimated proportion in group two
q_2 = 1- p_2
d = the desired precision in detecting a difference

Your project will need to make some assumptions in terms of the estimated proportions in each group. Refer to existing data (MOH data, data from other PVOs or NGOs working in the same region, DHS studies) to come up with reasonable estimates.

THE TAKE-HOME MESSAGE: WHEN IS A SAMPLE BIG ENOUGH?

It is important to balance the desire for precision with the amount of effort that has to be invested to conduct more interviews. Each project that is interested in increasing its sample size should ask, "How much would we gain from the extra effort to collect additional data?" Unless your project is interested in making comparisons—either across time, between control and intervention communities, or between different segments of the target population—a sample size of 300 is probably adequate. Remember: The main purpose of the KPC is to provide descriptive data for decisionmaking. It is not intended for hypothesis testing or other forms of analytic research.

Guiding Principle #3: Include what you can use

Parallel Sampling

More projects are targeting different population groups—not just young children and their mothers—in their interventions. For example, husbands or mothers-in-law might be the targets of special Behavior Change Communication (BCC) messages. Adolescents might be the focus of an HIV/AIDS campaign. Nonpregnant women of reproductive age (ages 15–49) might be important targets of activities promoting child spacing, HIV/AIDS, or women's nutrition awareness.

It is important to collect baseline information on all of your population targets. Sometimes, it is possible to do this through a KPC study, using a technique called *parallel sampling*. Parallel sampling is an easy and efficient way to collect information on different target groups within the same study. With parallel sampling, one randomly selected household serves as the basis for more than one interview

with different types of respondents. Parallel sampling can be used with different sampling methods, not just cluster sampling.

By including different types of respondents in a study, your project can maximize the usefulness of the KPC as a planning tool. In addition, parallel sampling can save time and money that would otherwise be needed to conduct separate studies with each population group of interest.

Below are some reasons to use parallel sampling.

- To collect special information from certain population groups, for example:
  - ⇨ Women of reproductive age (child spacing; HIV/AIDS; health contacts/sources of information)
  - ⇨ Husbands (knowledge of maternal, newborn, and child danger signs; treatment of childhood illnesses; child spacing; HIV/AIDS)
  - ⇨ Mothers-in-law (knowledge of maternal, newborn, and child danger signs; treatment of childhood illnesses; health contacts/sources of information)
  - ⇨ Non-maternal caregivers (knowledge of maternal, newborn, and child danger signs; treatment of childhood illnesses; health contacts/sources of information; infant/child feeding; hand washing)
  - ⇨ Adolescents (HIV/AIDS and other STIs)

- To improve the precision and accuracy of certain types of data

As mentioned before, many KPC indicators refer to subsamples. Because precision is related to sample size, one can expect that an indicator based on a subsample will be less precise than an indicator based on the entire sample.

A project can use parallel sampling to make sure that there are adequate numbers of cases in each subsample. For example, the project might decide to divide mothers into two groups: mothers of children 0–11 months and mothers of children 12–23 months. In each cluster, interviewers could conduct 10 interviews for each group of mothers. This would provide your project with a greater number of cases for subsamples based on age.

VERY IMPORTANT: There are usually questions that are asked of all mothers, regardless of the child's age (for example, treatment and care-seeking practices for sick children; hand-washing behavior; child spacing). If this is the case, **you do not want to over represent the practices and behaviors of a particular household in your study by interviewing more than one mother in the same household. Interviewing two mothers who live in the same household increases the design effect because those mothers probably share similar behaviors and practices. This would result in a biased assessment.

There is another benefit to dividing mothers into two groups based on the current age of the child. A person is less likely to accurately recall details of events that occurred further back in time. This is known as *recall bias*. Due to this potential bias, some projects limit questions about pregnancy, delivery, the postpartum period, and immediate newborn care to mothers of children 0–11 months only. Because those mothers were pregnant in the recent past (within the last year), there is a better chance of them remembering information about the pregnancy.

AN IMPORTANT RECOMMENDATION WHEN USING PARALLEL SAMPLING

To avoid confusion, your project should create different questionnaires for each target group. For any given target group, the questionnaire should only include questions that you will ask of respondents in that group. For example, you would not include questions on breastfeeding practices or care-seeking practices for sick children in a questionnaire for adolescent males.

Clearly label which questionnaire is for each target group in large letters on the first page of each questionnaire. This will make it easier for interviewers to quickly distinguish between the different forms when they are in the field.

- To reduce the length of interviews

Because interviewers will direct some questions to certain types of respondents and omit them for others, the time spent interviewing each respondent can be drastically reduced.

The following page presents a scenario on using parallel sampling.

AN EXAMPLE OF PARALLEL SAMPLING

Your project is interested in the following population groups:

1. Nonpregnant women 15–49
2. Mothers of children 0–11 months
3. Mothers of children 12–23 months
4. Men 15–49.

The KPC coordinating team has decided to use parallel sampling. In each cluster, it wants to conduct 10 interviews for each group, and it has designed a separate questionnaire for each. The two questionnaires targeting men and women of reproductive age only contain questions on knowledge of child spacing methods, HIV/AIDS, and sexually transmitted diseases.

One of your interviewers visits a household where there are two sisters and their sister-in-law living in the same household. Each woman has a husband who is between 15 and 49 and also lives in the household. One of the women is 19 and has no children. Her sister is 21 and has a child who is 18 months. Their sister-in-law has a baby who is 3 weeks. None of the women are currently pregnant. What should the interviewer do?

The interviewer can administer three questionnaires in that household:

1. One questionnaire for women 15–49
2. One maternal questionnaire (EITHER the one for mothers of children 0–11 months OR the one for mothers of children 12–23 months)
3. One questionnaire for men 15–49

Any of the women can answer the questionnaire designed for women 15–49. However, the interviewer will need to randomly select (see Table 5.2) one of the two mothers in the household, and then administer the correct questionnaire based on the age of her child. Can the same woman answer questions from both the 15–49 questionnaire and the maternal questionnaire? Yes, if the two questionnaires do not contain the same questions, or the project will not be aggregating the data from nonpregnant women and mothers of young children. In those instances, it would be okay to administer the two questionnaires to the same woman. In terms of the men's questionnaire, the interviewer can randomly select one of the husbands for the interview.

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

Guiding Principle #4: Use what you know

Stratified Sampling

For program-wide estimates, cluster sampling is a good option if your project knows very little about the target population and it wants to invest modestly in the data collection process. However, if your project has information on population characteristics that might influence maternal and child health outcomes, you should consider using that information when designing a KPC. This can be achieved through stratified sampling. The concept of stratified sampling was introduced earlier, when we discussed ways to make sure that the KPC sample in the make-believe District X included respondents from remote communities.

There is always a chance that a sample is not fully representative of the general population. This is especially true if the general population contains a small number of people from a particular group (for example, a certain religious or ethnic group). If there is only a small chance that people from a small group will be sampled, there is a possibility that a simple random or cluster sample might not include any (or very few) respondents who are members of that particular group. If your sample does not include key population groups, or if the proportion of certain groups in your sample differs from their proportion in the larger population, there is a chance that your survey's estimates will be inaccurate. In other words, the estimates based on the sample differ from the actual values that exist in the general population.

With both simple random sampling and cluster sampling, there is a chance that certain groups will not be included in the survey. What can a project do to make sure that the sample is representative in terms of a certain characteristic (e.g., urban-rural residence, ethnic group)? It can use stratified sampling. Using stratified sampling, you can make sure that important population groups are represented in the sample. In addition, by sampling adequate numbers of people in each group, a project can explore differences between groups. With stratified sampling, you first subdivide (stratify) the population into more homogeneous strata, according to a characteristic of interest. You then collect random samples within each subdivision (stratum).

The following are common ways to stratify a sample:

- Program management units such as health facility catchment areas or supervision areas
- By place of residence
- By language/ethnic group
- By religious group
- By socioeconomic status (There may not be existing socioeconomic data for your program area, but use what you know—for example, information on type of

dwelling, access to electricity, ownership of certain durable goods such as bicycles, televisions).

How is stratified sampling an improvement over simple random and cluster sampling? In addition to ensuring that certain groups are represented in the sample, stratified sampling can also be used to reduce the sampling error (resulting in estimates that are more precise). As mentioned above, stratified sampling involves dividing the program area into strata (smaller subdivisions). Each stratum is more homogeneous in terms of a particular trait or set of traits. In other words, it contains individuals with like characteristics (the same language, religion, or socioeconomic level). With stratified sampling, the sampling error is calculated based on variation within stratum (unlike cluster sampling, for which you compute the sampling error based on variation between clusters). Because of this fact, estimates from a stratified sample are more precise.

With stratified sampling, a project has the option of either a) sampling a certain proportion of cases from each stratum, or b) sampling a set number of cases from each stratum. With the first option, sampling a set proportion in each group results in a sample distribution that mirrors the population distribution. For example, assume that the population in your project area can be divided into four ethnic groups, and that your project has decided to sample 5 percent of the members in each group. As seen in Table 5.7, the sampling fraction (the number of cases in the sample divided by the number of cases in the population) is the same in each stratum. The sample distribution is the same as the population distribution.

Table 5.7
Sampling 5 Percent of the Population in Each Stratum (Ethnic Group)

Ethnic Group	In the Population		Sampling Fraction	In the Sample	
	Number	*Percentage*		*Number*	*Percentage*
A	2,000	17%	.05	100	17%
B	1,000	8%	.05	50	8%
C	6,000	50%	.05	300	50%
D	3,000	25%	.05	150	25%
TOTAL	12,000	100%	.05	600	100%

Suppose your project is interested in making comparisons between the four ethnic groups. As seen above, the sample only contains 50 cases from Ethnic Group B. At the other extreme, there are 300 cases in Ethnic Group C. Your project would like to make some group comparisons for certain indicators, but there would not be enough cases in each group, according to the approach in Table 5.7. To get enough respondents in each group, the project could sample a set number of cases in each stratum. How many cases would it need to sample? A staff member of your local partner organization is very good at sampling. She volunteers to meet with the KPC coordinating team to do some sample size calculations to determine an appropriate sample size, given the amount of difference you want to be able to detect between groups. She and the KPC coordinating team come up with 150 respondents in each group.

Table 5.8
Sampling 150 Individuals in Each Stratum (Ethnic Group)

Ethnic Group	In the Population		Sampling Fraction	In the Sample	
	Number	*Percentage*		*Number*	*Percentage*
A	2,000	17%	.075	150	25%
B	1,000	8%	.150	150	25%
C	6,000	50%	.025	150	25%
D	3,000	25%	.050	150	25%
TOTAL	12,000	100%	.050	600	100%

As seen in Table 5.8, the sampling fraction varies across groups. For example, only 2.5 percent of people in Ethnic Group C were sampled, whereas as 15 percent of people in Ethnic Group B were sampled. However, you now have an equal number of cases in each group (notice that each group contributes 25 percent to the total sample size), which will allow you to make some preliminary comparisons between groups for select indicators.

There is one problem: Your sample distribution no longer mirrors your population distribution. There is a way to aggregate (combine) data across all strata to get estimates that apply to the program area as a whole. This is done by weighting the data from each stratum based on the population distribution.

To illustrate how to weight data from a stratified sample whose distribution is different from the distribution in the general population, let's use the following *Rapid CATCH* indicator:

Percentage of mothers of children 0–23 months who know at least two signs of childhood illness indicating the need for treatment

Assume that 10 percent of mothers of Ethnicity A, 20 percent of mothers of Ethnicity B, 33 percent of mothers of Ethnicity C, and 40 percent of mothers of Ethnicity D know at least two signs. Your project wants to calculate an overall estimate for the entire program area. In Table 5.9, the overall estimate is calculated based on a sample in which the same proportion of cases was sampled from each group. In Table 5.10, the overall estimate is based on a sample in which the same number of cases was sampled from each group. The same data from Table 5.10 are presented in Table 5.11, but this time the data are weighted by the proportion of individuals in each ethnic group in the general population. As you will see on the next page, the overall estimates from Tables 5.9 and 5.11 are the same (30 percent). Note, however, that the unweighted estimate in Table 5.10 is not that different (26 percent).

Table 5.9
SCENARIO 1: SAMPLE DISTRIBUTION EQUALS POPULATION DISTRIBUTION
Stratified Sample Estimates of Maternal Knowledge of Child Danger Signs
Overall sample estimate=30%

Ethnic Group	No. of Cases Knowing at Least Two Signs	Total Number of Cases	Number in the Population	Percentage Knowing At Least Two Signs
A	10	100	2,000	10/100= .10
B	10	50	1,000	10/50= .20
C	100	300	6,000	100/300= .33
D	60	150	3,000	60/150= .40
TOTAL	180	600	12,000	**180/600=.30, or 30%**

Table 5.10
SCENARIO 2: SAMPLE'S DISTRIBUTION DIFFERS FROM POPULATION'S
Stratified Sample Estimate of Maternal Knowledge of Child Danger Signs
Overall sample estimate (UNWEIGHTED)= 26%

Ethnic Group	No. of Cases Knowing at Least Two Signs	Total Number of Cases	Percentage Knowing At Least Two Signs
A	15	150	15/150= .10
B	30	150	30/150= .20
C	50	150	50/150= .33
D	60	150	60/150= .40
TOTAL	150	600	**150/600=.26, or 26%**

Table 5.11
SCENARIO 3: SAMPLE DISTRIBUTION DIFFERS FROM POPULATION'S
BUT DATA ARE WEIGHTED TO CALCULATE RATE FOR ENTIRE SAMPLE
Stratified Sample Estimates of Maternal Knowledge of Child Danger Signs
Overall sample estimate (WEIGHTED)=30%

Ethnic Group	No. of Cases Knowing at Least Two Signs	Total Number of Cases	Percentage Knowing At Least Two Signs		
			(1) UNWEIGHTED DATA	(2) WEIGHTS (Pop. Distribution)	(3) WEIGHTED DATA columns (1) x(2)
A	15	150	15/150= .10	.17	.10 x .17= .017
B	30	150	30/150= .20	.08	.20 x .08= .016
C	50	150	50/150= .33	.50	.33 x .50= .165
D	60	150	60/150= .40	.25	.40 x .25= .100
TOTAL	150	600	**150/600=.26**	**1.0**	**.298, or 30%**

REMEMBER TO DRAW RANDOM SAMPLES WITHIN STRATA

There are many benefits of stratified sampling over simple random or cluster sampling. However, the division of the population into homogeneous strata is only the first step. The advantages of stratified sampling are dependent on whether you adhere to the "Make random the standard" principle. Therefore, remember to select samples randomly within each stratum.

This is particularly important when you are sampling a small proportion of cases from a group (such as Ethnic Group B in Table 5.8).

What are the next steps if you discover that there are differences between strata? It helps to investigate the problem further. During the KPC Analysis Workshop (held after data collection is completed), explore reasons why those differences might exist. Also, explore whether your project should have different objectives (set different targets) for each stratum, or whether the project needs to alter the types of activities or the amount of resources, depending on the stratum.

LQAS—A Special Form of Stratified Sampling

Recently, CS projects have become interested in a special form of stratified sampling called Lot Quality Assurance Sampling (LQAS). There are two key differences between LQAS and regular stratified sampling:

- When a project divides the target population into socioeconomic characteristics such as ethnicity, religion, or socioeconomic status, it is acknowledging that there might be certain confounders (external factors) that affect maternal and child health outcomes. However, those factors are not very responsive to short-term intervention. More specifically, it is very hard to change long-standing and deeply rooted cultural or religious practices within a 5-year period—the time frame now allotted to CS projects to implement activities and demonstrate program effectiveness. Instead, projects focus on factors that they can more likely change, such as health worker performance. The concepts of "lots" and "production units" (see Table 5.12 for definitions) are important in LQAS. With LQAS, you divide the population into service delivery areas or program management units. These subdivisions serve as lots (strata) in LQAS. Common strata or lots for LQAS are health facility catchment areas or project or MOH supervision areas. The production unit is usually a health worker or a team of health workers and possibly their clients.
- Stratified sampling techniques such as LQAS allow projects to draw comparisons between subdivisions of a population. However, the main objective of LQAS is not to obtain individual estimates from those subdivisions, but rather to base program management decisions using a *binomial* principle. Binomial means that there are only two possible answers or outcomes (for example, yes or no; high or low). In other words, with LQAS, you are not determining the level of coverage in each subdivision. Instead, you are determining whether coverage in each subdivision is one of two things: a) at or above expectation, or b) below expectation.

When deciding on lots, it helps to stratify the population in a way that is not only meaningful to the project but also to the MOH and other stakeholders. Projects are encouraged to build assessments using LQAS into an on-going system of monitoring and supportive supervision. Staff from the MOH and local NGOs are more likely to adopt and maintain this information-gathering strategy if it produces data that are relevant to their activities and needs.

Table 5.12

Common LQAS Concepts

- *Lot:* In health applications of LQAS, lots tend to be supervision areas or catchment areas of health facilities. Each lot usually consists of several villages or communities.
- *Production unit:* From a health perspective, the production unit is usually a health worker or a team of health workers within each lot. In other words, it is an implementation unit in your project.
- *Decision rule*: The basis for judging whether a lot has coverage that is at or above expectation versus below expectation. The decision rule is a predetermined number out of a set of observations, and it is based on *lower* and *upper* thresholds.
- *Lower threshold:* The cut-point below which a lot is classified as "below expectation."
- *Upper Threshold:* The minimum standard that must be met for a lot to be classified as "acceptable" (at or above expectation) in terms of coverage.

A few words about sample size

The number of observations (sample size) within each lot and the decision rule are based upon what is statistically acceptable for the health manager or supervisor. Ideally, the sample size should be large enough that the manager has a high probability of identifying lots that are at or above the upper threshold and a high probability of identifying lots that are at or below the lower threshold. Project managers should identify the smallest possible sample size to keep the risk of misclassification below 10 percent for all indicators of interest. Samples of size 10–19 generally satisfy this criterion. **However, a sample size of 19 is often used because this sample size gives managers the flexibility to change upper thresholds after data have already been collected, without running the risk of having an insufficient amount of data to test those thresholds**.

This field guide does not present step-by-step instructions on how to implement a KPC survey using LQAS. Instead, it provides a general overview of the sampling strategy and presents key issues to consider. The documents listed at the end by Valadez (1991) and Valadez et al. (2000) provide detailed information on how to use LQAS.

LQAS, which in simple terms is just random sampling within service delivery areas (lots), is often compared and contrasted with the 30-cluster methodology.

Table 5.13 compares the two sampling methods.

Table 5.13
Cluster Sampling versus LQAS: Issues to Consider

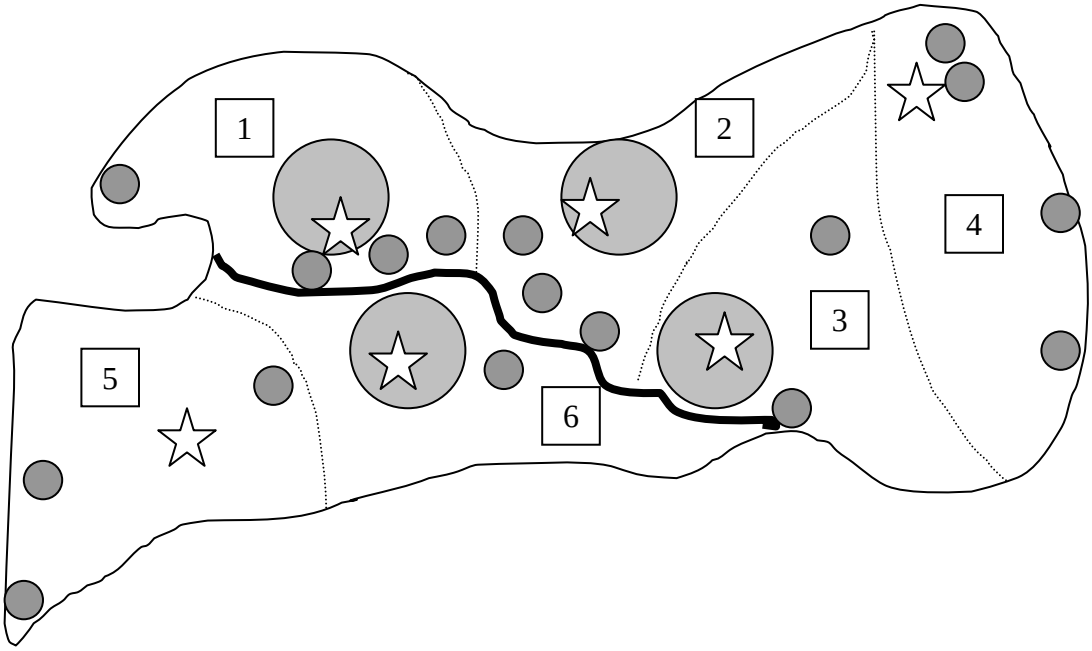
ISSUE	SAMPLING METHOD	
	CLUSTER	LQAS
Cost	• Usually cheaper than simple random sampling; it might also be cheaper than LQAS if data collection is not localized (i.e., a central team of interviewers is responsible for collecting data in all subdivisions).	• Can be more expensive than cluster sampling if a central team of interviewers is responsible for collecting data in all subdivisions. • Building LQAS into an on-going system of supervision can reduce costs. Certain personnel can be responsible for particular subdivisions within the project area in conjunction with routine support activities. In doing so, it is possible to survey an entire program area at a lower cost than if the project used cluster sampling.
Usefulness as a Method at Baseline	• Fast and efficient way of obtaining program-wide information, particularly when a PVO is new to a region and does not yet know how to divide the population into subdivisions that are programmatically meaningful. • Cluster sampling can be used to give an idea of the scope of a problem for the entire program area. It does not shed light on which subdivisions within the area might require additional resources for the project to achieve its objectives. If there are both high and low coverage areas in the program area, these differences would be missed if cluster sampling were used.	• Can highlight which subdivisions within the project area have levels of coverage that are at or above average versus those that are below average. • LQAS can be useful at baseline if there are 1) clearly defined (and programmatically meaningful) program management units (such as MOH supervision areas or health facility catchment areas), and 2) "production units" (such as teams of community health outreach workers). • If certain areas are classified as above average, and others as below average in terms of coverage, a project might decide to set different objectives for certain areas. Being able to identify areas of greatest need can also help in targeting resources more effectively.
Usefulness as a Method for Project Monitoring	• May be used to provide program-wide estimates of progress toward targeted objectives • Not very useful for monitoring purposes because it does not identify high and low performance areas within a project.	• Provides management information at the local level by determining which areas are at or above a certain threshold level versus those that are below that threshold level. • Although LQAS can identify *which* subdivisions are performing below expectation, it does not shed light on *why* they are performing below expectation. A project will need to further investigate process issues that might explain why low coverage or inadequate service delivery exists.
Usefulness as a Method at the End of a Project	• Allows a project to assess whether or not program objectives have been met for the entire project area.	• If different objectives were set for different subdivisions, LQAS will not tell you whether objectives were met for each subdivision. • Unless there are plans for a follow-on project, identifying which areas are below or above average will not be very useful at the end of a project.

Table 5.13 (continued)
Cluster Sampling versus LQAS: Issues to Consider

ISSUE	SAMPLING METHOD	
	CLUSTER	**LQAS**
Precision of Aggregate (Program-Wide) Estimates	• Usually less precise than estimates that would be obtained from a simple random sample or lot quality assurance sample of the same size. • The precision of estimates based on cluster sampling depends greatly on the extent to which a cluster is homogeneous relative to the population (design effect).	• Can yield aggregate survey estimates that are more precise than estimates from a cluster sample of the **same size.** • In other words, if you compared estimates from a LQAS aggregate sample of 300 with estimates from a cluster sample of 300, the LQAS estimates would probably be more precise (have narrower confidence intervals). • It is important to note that if you desire a certain level of precision in your aggregate estimates, this should be taken into account when determining how many interviews to conduct in each lot (subdivision). If the total number of interviews across all lots is less than 300, your aggregate-level estimate might not be very precise, even though you are sampling from every stratum.
Local Estimates	• Only provides estimates for an entire program area. • A project can stratify the population into meaningful subdivisions (strata), and then select enough cases in each stratum to yield an estimate for each stratum. You will need to do some sample-size calculations to determine how many cases are needed in each stratum in order to achieve a certain level of precision for each estimate. Keep in mind that although this is an option, there are money and time implications in collecting additional data.	• Not the purpose of LQAS; does not provide coverage estimates for each subdivision of a program area, only assesses whether each subdivision is above or below a particular level of coverage considered as acceptable.
Population Density	• In widely dispersed populations (people live far apart from one another), it is an efficient way of gathering data because it can reduce the time and money spent traveling between interviews.	• In widely dispersed populations, it can be time-consuming and expensive because interviewers will probably have to travel great distances between each randomly selected sampling point.
Information Gatherers	• Usually hire teams of interviewers to conduct interviews in clusters selected from the broader population. • Quality-control measures should be in place to ensure that households and respondents are selected properly. Interviewers should receive supportive supervision from the field supervisors during the data collection process.	• Can use certain personnel for particular subdivisions. In some applications of LQAS, supervisors are responsible for collecting the data in their supervision area. Although this provides a wonderful way for a manager/supervisor to self-assess the performance in his or her own area, projects should also consider the trade-offs in terms of quality control. Supervisors should be trained to be objective, so as not to introduce a bias in collecting information from the same communities where they work. • To avoid bias, each supervisor could collect data in another supervisor's area. Keep in mind, however, that this might impact the cost-effectiveness of using LQAS.

As an example of how a project could use LQAS, let's go back to the map of make-believe District X. In the map below, dashed lines have been added that divide the District into six health facility catchment areas.

Map of District X, Divided into Six Health Facility Catchment Areas



Suppose your project would like to use LQAS during its mid-term evaluation and at six-month intervals until the end of the project. Rather than conduct a full-scale KPC survey at midterm, the project decides to use 12 questions from the baseline KPC questionnaire. Those 12 questions are linked to five monitoring indicators.

The project has designated the six catchment areas as the lots for its periodic LQAS assessments. In each LQAS assessment, the project wants to determine which of those lots are on track in terms of meeting the program's objectives, and which lots require more focused attention.

The project decides to conduct 19 interviews in each lot. This sample size will allow the project to make an informed decision (whether the estimated level is at or above expectation versus below expectation) about each indicator for each lot, while limiting the risk of misclassifying a lot to 10 percent.

REMINDER

When using LQAS, the project's main objective is to classify lots into one of two groups:

1. Coverage at or above expectation
2. Coverage below expectation

The objective behind LQAS is not to get coverage estimates for each lot. This is one major difference between LQAS and other forms of stratified sampling. With LQAS, you collect information on only a small number of cases in each lot. (It is common to conduct **19** interviews in each lot.) Estimates based on such a small number of cases will have extremely wide confidence intervals, and therefore are not very meaningful. As a result, only combine data across lots to get estimates for the entire program area. These program-wide estimates are analogous to what you get with a cluster survey. With good quality control of data collection, estimates based on an LQAS sample can be more precise than estimates based on a cluster sample of the same size.

A SPECIAL NOTE ABOUT SUB-SAMPLES

Some projects that used LQAS in the past have mentioned that they encountered problems with sub-samples, particularly when assessing treatment and care-seeking practices for sick children. Some projects conduct 19 interviews in each lot, and then find that they do not have enough sick children in the sample to assess care-seeking practices. For example, there might be a small number of children who had pneumonia in the two weeks before the survey. Within that group, an even smaller number of children might have received treatment for the illness. How can a project make decisions when there is little or no data? There are two solutions:

- The project can decide to aggregate data across lots and calculate a program-wide estimate, rather than make assessments for each lot.
- If the project is interested in making assessments regarding sick children in each lot, it will need to randomly sample additional sick children. With LQAS, there are 19 randomly selected sampling points in a given lot. This might not necessarily correspond to the number of interviews conducted in each lot. For example, if your project uses parallel sampling to collect information on mothers of children 0–11 months, mothers of children 12–23 months, and nonpregnant women of reproductive age, it will need to conduct 19 interviews for each group. If your project is concerned about having too few sick children in the sample, it should plan to randomly select additional sampling points in each lot (that is, beyond the 19 sampling points that it would normally select) until it finds enough sick children in the lot. Your project will need to decide on that number. If an interviewer visits one randomly selected household and finds a sick child, he or she should not visit neighboring households to find additional sick children. If that is done, he or she is selecting a cluster sample of sick children! Instead, the interviewer should keep randomly selecting different sampling points (households) until he or she finds the desired number of sick children.

Remember, do not take the easy way out and sample more than one sick child within the same household. Nor should you go to a neighboring household. You have to randomly select a new household to avoid introducing a bias into the study.

Guiding Principle #5: The sky is not the limit

A project should try to maximize the usefulness of its KPC survey, and it should aim to incorporate any information that will help to plan and manage the project. However, it is important to realize that the KPC cannot provide answers to all of your questions. Certain topics are beyond the scope of a KPC survey. Also, other tools and information-gathering methods might be more appropriate to explore certain phenomena. When choosing whether to include certain types of questions or respondents in the survey, it is important to consider the following two questions:

- Can we be confident in the quality of the information that we are collecting?
- Is there a better way (tool, methodology) to collect this information?

Information that does not relate directly to program management, or that requires a different methodology to be accurate and reliable, should not be collected in your KPC survey.

SUMMARY OF SIMPLE RANDOM, CLUSTER, AND STRATIFIED SAMPLING

Figure 5.1 and Table 5.14 summarize the main sampling methods discussed in this section of the guide.

Figure 5.1
Differences in the Precision of Estimates Based on Cluster, Simple Random, and Stratified Sampling

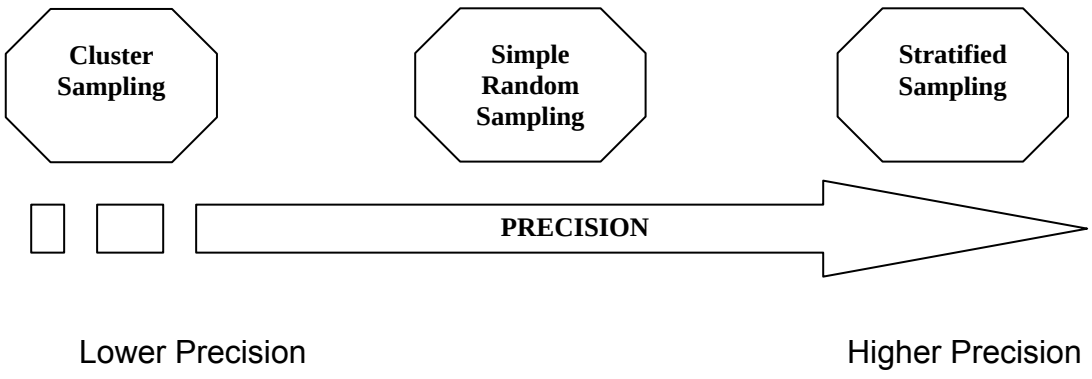


Table 5.14
A Comparison of Simple Random, Cluster, and Stratified Sampling

	SIMPLE RANDOM SAMPLING	CLUSTER SAMPLING	STRATIFIED SAMPLING
General Description	Sampling units are selected *randomly* from a *sampling frame*. Every element in the sampling frame has an equal chance of being included in the sample.	Clusters (rather than individuals) are randomly selected. Cluster selection is based on probability proportional to size so that large communities have a greater chance of having clusters than smaller communities. Within each cluster, several individuals are randomly selected to reach the desired sample size.	The target population is first divided into *strata*, then a random sample is selected from each stratum, ensuring that individuals from every stratum are represented in the sample.
Design Effect	None. SRS is the standard against which all other sampling designs are compared when estimating the design effect.	A bias that results from randomly sampling clusters, rather than individuals. Usually between 1.5 and 2. To be conservative, assume a design effect of 2. That means that there must be twice as many respondents in a cluster sample compared to a SRS, in order to compensate for the design effect.	Stratified sampling is more precise than SRS, provided that the strata are homogeneous and sampling units (households, individuals) within each strata are selected randomly.
Sample Size	Equal to 96 for a cross-sectional study with no comparison groups.	At least two times larger than a SRS to compensate for the design effect.	Can be smaller than both SRS and cluster sampling (for example, when using LQAS).

Table 5.13 (continued)

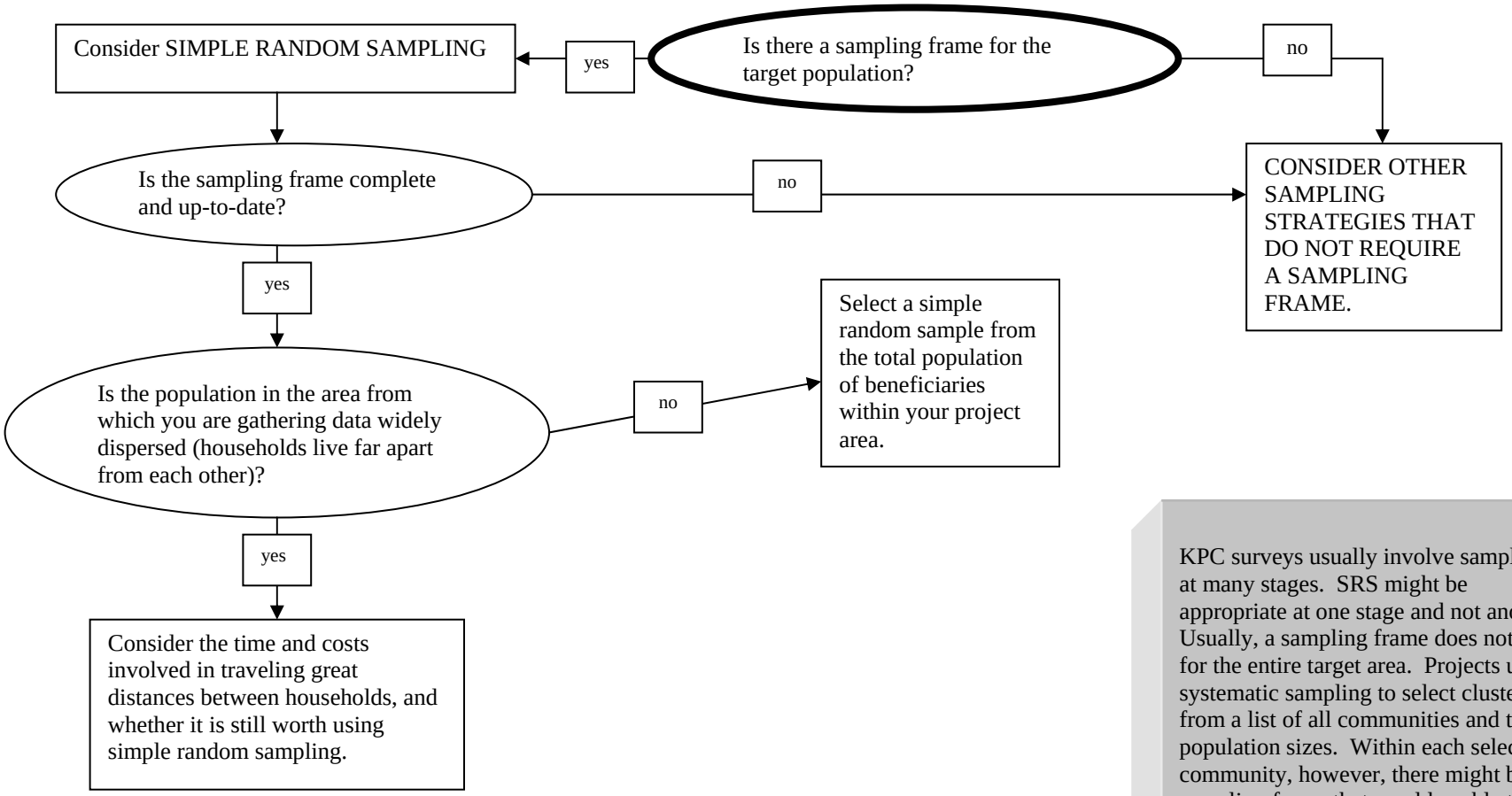
	SIMPLE RANDOM SAMPLING	CLUSTER SAMPLING	STRATIFIED SAMPLING
Advantages	Straightforward to execute if a sampling frame is available.	• Saves time by surveying several individuals within a cluster rather than repeating the process of randomly selecting each individual to reach the required sample size for the survey (as done with SRS). • Cheaper to execute than SRS. • Maximizes variability in the sample while minimizing study costs. • Efficient way of sampling in dispersed populations. • A good option when little is known about the study context.	• Ensures that groups of interest are covered in the survey. • The standard error of each estimate is based on the variation within homo-geneous strata. Because of this, stratified sampling yields estimates that are more precise than estimates from SRS or cluster sampling. • There is the option to disproportionate-ly sample within strata to get enough cases from each strata in your sample (important if you want to make comparisons between strata).

Table 5.13 (continued)

	SIMPLE RANDOM SAMPLING	CLUSTER SAMPLING	STRATIFIED SAMPLING
Disadvantages	• Requires a complete and up-to-date sampling frame, which is usually not available for large populations. • All groups of interest might not be included in the sample. • Expensive and inefficient in large or dispersed populations.	• Cannot get estimates for subdivisions within the program area. • A bias (the design effect) is introduced by interviewing persons in clusters rather than as randomly selected individuals. • Lower precision than SRS or stratified sampling because standard errors (measures of precision) are based on variation between homogeneous clusters.	

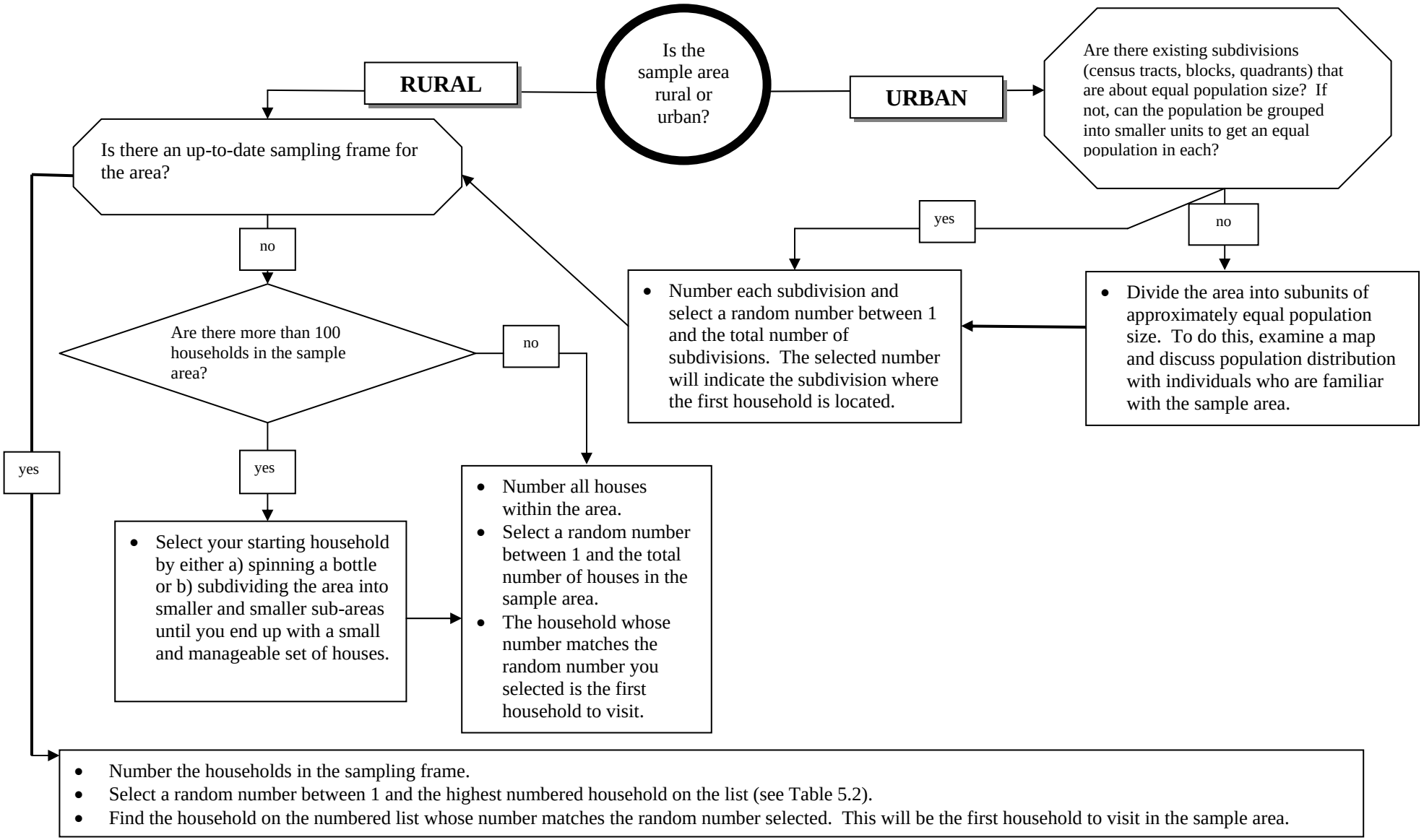
IS SIMPLE RANDOM SAMPLING APPROPRIATE FOR YOUR PROJECT?

Start Here

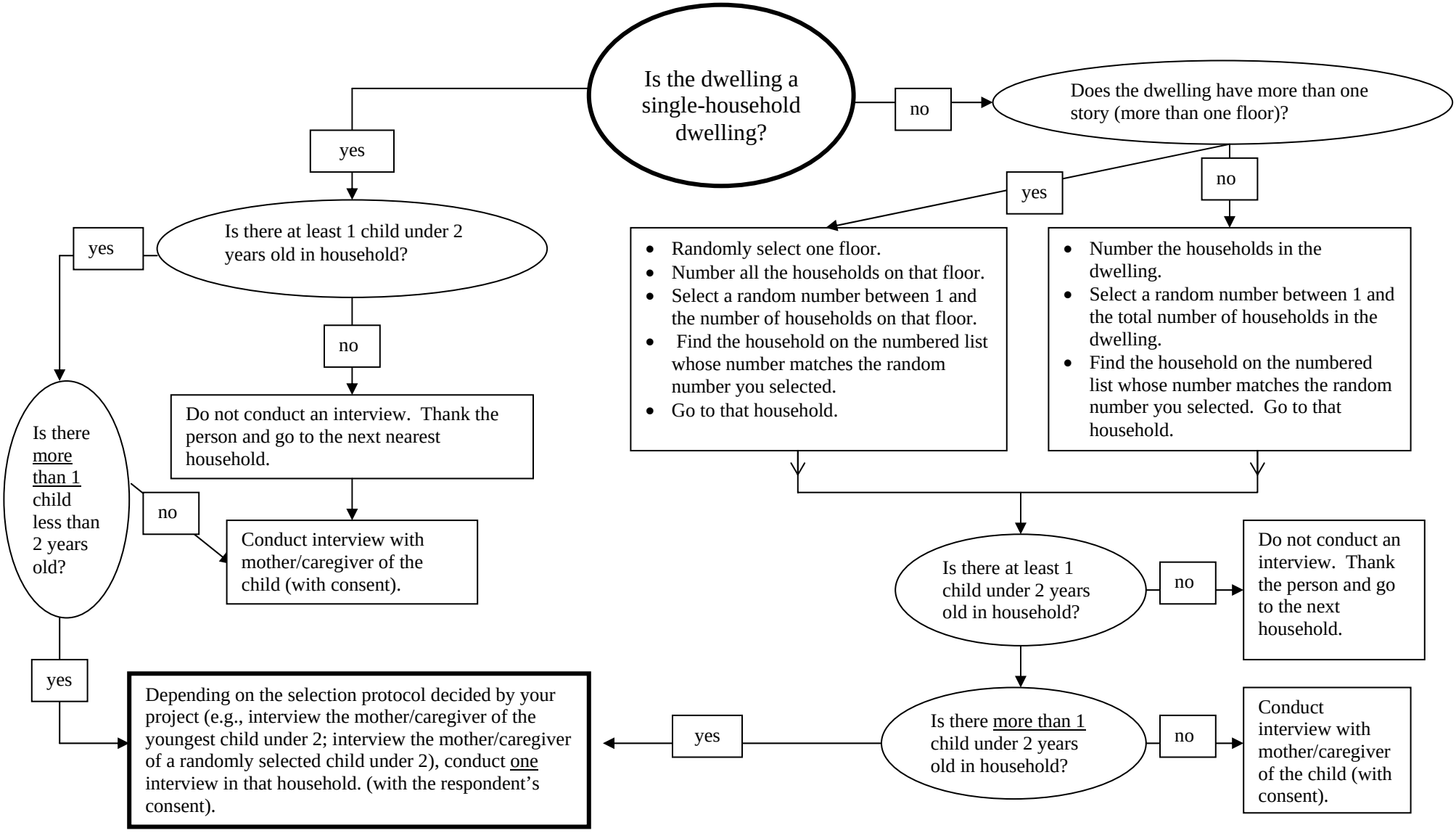


KPC surveys usually involve sampling at many stages. SRS might be appropriate at one stage and not another. Usually, a sampling frame does not exist for the entire target area. Projects use systematic sampling to select clusters from a list of all communities and their population sizes. Within each selected community, however, there might be a sampling frame that would enable the project to use SRS to select households and respondents.

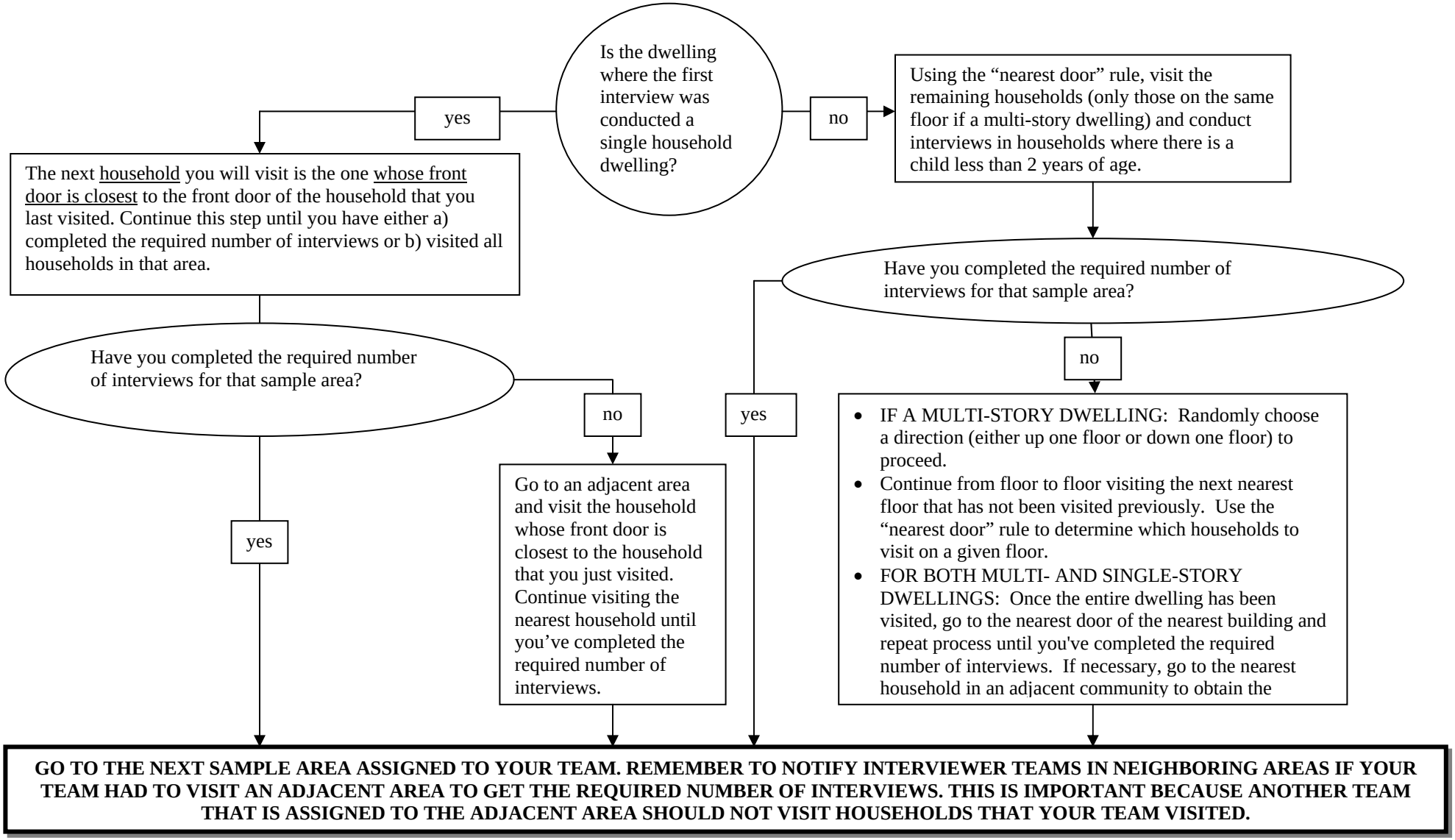
SELECTING THE FIRST HOUSEHOLD IN A SAMPLE AREA



CHOOSING RESPONDENTS BASED ON THE TYPE OF DWELLING



CONDUCTING THE REMAINING INTERVIEWS IN THE SAMPLE AREA



USEFUL RESOURCES

Aday, L. 1996. *Designing and conducting health surveys: A comprehensive guide*. San Francisco: Jossey-Bass Publishers.

Centers for Disease Control. Epi Info Software and Documentation. **Available on the CDC Web site (www.cdc.gov/epiinfo).**

Habicht, J., C. Victora, and J. Vaughan. 1999. Evaluation designs for adequacy, plausibility, and probability of public health programme performance and impact. *International Journal of Epidemiology* 28, 10–18.

Henderson, R., and T. Sundaresan. 1982. Cluster sampling to assess immunization coverage: A review of experience with a simplified sampling method. *Bulletin of the World Health Organization* 60(2), 253–260.

Sarriot, E., P. Winch, W. Weiss, and J. Wagman. 1999. *Methodological and sampling Issues for KPC surveys*. **Available at CSTS Web site (www.childsurvival.com) under KPC2000+.**

Valadez, J., W. Weiss, L. Seims, R. Davis, and C. Leburg. 2000. (DRAFT) *Using LQAS for assessing field programs in community health in developing countries. A trainer's guide for baseline surveys and regular monitoring*. **Available from NGO Networks for Health.**

Valadez, J. 1991. *Assessing child survival programs in developing countries: Testing lot quality assurance sampling*. Boston: Harvard University Press.

World Health Organization. 1991. *Training for mid level managers: The EPI cluster survey*. WHO/EPI/MLM/91.10.

6. RECRUIT AND TRAIN SUPERVISORS AND INTERVIEWERS

Recruit Field Personnel

Selecting Supervisors

It is helpful to develop a list of desired qualifications before selecting supervisors. For example:

- Must be literate—can read and write
- Should have some field or health experience
- Must speak the major local language(s)
- Must be available for the full survey and related activities
- Must be organized sufficiently to supervise interviewers and maintain data integrity
- Must be a dependable worker
- Should know the general area
- Must be physically fit
- Must be able to accurately audit completed questionnaires
- Should have supervisory experience

The following types of individuals are potential candidates:

- Project and partner staff
- Midwives
- Nurses
- Students from health and sanitation schools
- University students
- Medical Auxiliary students
- Staff from partner PVOs/NGOs

- MOH staff
- Allied health sector personnel—hospital or clinic staff, etc.
- Community based workers and/or volunteers
- Other individuals with previous interviewing experience in the project area

Selecting Interviewers

A similar list of qualifications can be developed to assist in the selection of interviewers. For example:

- Must be literate—can read and write well
- Must write legibly
- Have some field or health experience
- Must speak the household local language fluently
- Must be able to establish rapport easily
- Must work well as part of a team
- Must be available full-time for training, data collection, and tabulation activities
- Must be able to organize work and interview forms sufficiently to accurately record answers
- Must be dependable
- Must know the general survey area
- Must be physically fit

REMEMBER

When selecting interviewers, consider the impact of gender dynamics on the quality of an interview and therefore, the quality of survey data. In many societies, it is not acceptable for a woman to have contact with a male other than her husband or relatives. Even when such restrictions do not exist, some females might feel more comfortable talking to another woman than to a man. Therefore, consider using only female interviewers. Keep in mind, however, that if other groups are targeted in the survey—for example, husbands or adolescent males—male interviewers may be more appropriate for those population groups.

The ability to communicate in the local language or dialect is absolutely necessary for interviewers. An interesting exercise when recruiting interviewers is to ask groups of them to

role-play scenes of local life in the local language or dialect. Ask the interviewer candidates who are observing the role-plays to either write or verbally translate what they observed in the project language.

Training Field Personnel

Who Conducts the Training?

The training facilitator does not have to be the survey coordinator. In fact, projects are encouraged to select more than one person from the KPC coordinating team to facilitate at different times during the training. Different individuals have different strengths, and those who do not already have the capacity to train can be given the opportunity to develop such skills during the training process. Remember, it is important to build local capacity to conduct KPC surveys and to develop the skills required to be good facilitators.

How Long Should the Training Last?

Ideally, interviewer training should take place in 3–4 days, allowing for sufficient time to review all aspects of field implementation and the opportunity to get hands-on field experience conducting interviews. It is recommended that you hold a special one-day overview training with the field supervisors before conducting the interviewer training. During the overview session, the facilitator and supervisors can do the following:

- Review the purpose of the survey
- Discuss and agree upon roles of the supervisor and interviewer
- Review the questionnaire (item by item, paying attention to skip patterns and special instructions)
- Review proper interviewing and supervision techniques
- Practice an interview, receive feedback
- Review quality-control procedures in the field
- Discuss the role that supervisors can play during interviewer training.

In addition to the one-day overview, supervisors should attend the interviewer training session in its entirety.

What Should the Training Cover and How Should It Be Structured?

Although the nature of the training is left to the discretion of each project, a KPC training should review at least the following:

- Survey purpose

- Content and format of the adapted questionnaire
- Proper interviewing techniques
- Field procedures, including protocols for household and respondent selection
- Role plays
- Practice interviews and final field-testing of the questionnaire before beginning data collection.

There may be other topics that need to be discussed given the local context. In addition to instructional training, there should be an opportunity for interviewers to conduct practice interviews through role playing and in the field. Supervisors and training facilitators can assess and give feedback on each person's performance at that time. It is helpful for interviewers to practice the field procedure in its entirety—from going to a surveyed area, to randomly selecting the first household, identifying an eligible respondent within that household, conducting an interview, then selecting the next household. Supervisors should practice observing interviews, reviewing completed questionnaires for errors, and filling out quality-control checklists (see next chapter).

It is not necessary to discuss the sampling design or sample size calculations in great depth during the training sessions. However, if some interviewers or supervisors express an interest in learning more about those topics, then trainers can hold special sessions with individuals who want additional information on sampling.

During their one-day training, supervisors can help to refine the KPC questionnaire by conducting 1–2 practice interviews. Get their feedback on the questionnaire (format, length of the interview, skip patterns, wording of questions). Based on their input, further changes can be made to improve the questionnaire before reproducing it for the interviewer training. Field-testing can also be built into the interviewer training when interviewers are conducting practice interviews in the field. Interviewers might identify additional changes that should be made to the questionnaire before sending it off for final duplication.

Regardless of the types of activities that make up the interviewer training, facilitators should strive to create a trusting atmosphere that is conducive to team building and supportive supervision. The training is as much an opportunity for consensus building as it is for capacity building. Although there are certain topics that have to be covered in a KPC training workshop, the structure should be as flexible and participatory as possible. As part of team building and consensus building, it is helpful for the group to come up with certain principles or traits that they are committed to. For example, the group can come with two lists of traits that make a good supervisor and a good interviewer.

EXAMPLES OF PROPER INTERVIEWING TECHNIQUES THAT SHOULD BE DISCUSSED DURING SUPERVISOR/INTERVIEWER TRAINING

- Always be courteous, polite, respectful, and non-judgmental.
- Before interviewing a respondent, introduce yourself and state the name of the organization

you are working with and the general purpose of the survey.

- Maintain the confidentiality of the survey. If there are people around the respondent, ask them politely to leave. Be sure to do this in a way that is culturally appropriate. Explain to the respondent that he/she does not have to take part in the survey, that health services will not be withheld if he/she does not participate and that any information that he/she shares during the information will not be linked to his/her directly when examining the results from the survey. Gain the person's consent to be interviewed before asking questions.
- To begin with, ask each question exactly as it is written. When questions have to be repeated, use the phrases and minor wording changes that were agreed upon during the training.
- Ask questions in a respectful manner, do not imply that some answers are better than others.
- When an answer is unclear, ask the question again or ask it in a slightly different way, but be careful not to change the meaning or lead the respondent into a particular response. For example, suppose a caregiver mentions that the child was given "a special drink" while he or she had diarrhea. Do not ask a leading follow-up question such as "Do you mean that you used ORS?" Instead ask an open question like "What kind of special drink?" or "What was in the drink?"
- If an answer seems inconsistent with previous information given by the respondent, or if there is some reason to disbelieve an answer, try to discover the truth by asking him/her another question or asking a question in a slightly different way. However, do not be overly persistent. A person may change his/her answer just because persistent questioning suggests that the interviewer is dissatisfied with the initial answer.
- Ensure that translations of questions are not leading as some translations can prompt a particular answer.

PRACTICE, PRACTICE, PRACTICE

It is important to allocate a substantial amount of time for field practice and role-playing. This will allow interviewers the opportunity to master interviewing techniques, get comfortable with the layout of the questionnaire, and identify areas where they need additional practice before beginning the actual data collection. Field practice will also allow supervisors to practice observing interviews and assessing their quality as part of supportive supervision. At the end of each day of training, interviewers and supervisors should be instructed to practice conducting interviews in the evening with each other or with friends or family members. Make sure that practice interviews are not conducted in communities where interviews will be conducted.

REMEMBER

- Some individuals will need more attention and training than others. Training facilitators should assess the performance of each supervisor and interviewer daily and work to ensure that every person has mastered the necessary skills before data collection begins. In the spirit of team building, pair individuals who still need to develop certain skills with individuals who are strong in those areas. This process of identifying individuals whose skills complement one another also applies when survey team assignments are being made.
- Sometimes interviewers and supervisors have special insight into local customs and taboos. Facilitators and the coordinating team should be flexible in terms of further adapting the questions, methodology, or interviewing techniques to make them as appropriate for the local context as possible.

By the end of interviewer training, the coordinating team should do the following:

- Determine the team structure for each survey team
- Determine site assignments for each survey team
- Complete administrative and logistical preparations necessary to begin the conduct of the survey
- Assign transportation units, duplicate sufficient copies of the questionnaires, and distribute all supplies and materials needed to interviewers and supervisors
- Provide additional guidance to teams assigned to communities where either multiple survey sites (clusters, lots) are located or where urban areas were selected.

When choosing final interviewer teams and their assignments, a number of factors will need to be considered. For example:

- Observations during the training
- Local language proficiency
- Staff knowledge about individual workers

As stated earlier, try to assign teams whose member's skills complement each other. The survey coordinator and/or members of the coordinating team may opt to travel to the most difficult survey locations or to accompany interviewers and supervisors who performed least well during the training.

7. COLLECT THE DATA

The Survey Implementation Phase involves conducting KPC interviews in selected communities. **Quality control is critical to the data collection process.** The purpose of quality-control procedures is to maximize the performance of the interviewers and get the best possible data, given the circumstances of the local context.

Before sending survey teams into communities, the KPC coordinating team should have a clear-cut strategy for maintaining quality throughout the data collection process. It helps to create a Field Implementation Checklist, so that each survey team can take a daily inventory of all supplies and equipment before going into the field.

Supervisors should not conduct interviews. Their primary role is to support the interviewers, serving as the first point of contact when interviewers encounter problems in the field, and assessing and maintaining the quality of data collection. It is recommended that supervisors complete a Quality-Control Checklist for every interview they observe.

Members of the KPC coordinating team should visit the field periodically to assess data collection activities. However, while interviewers are collecting data, it is also a good time to finalize arrangements for tabulation, analysis, and dissemination.

Before sending interviewers out into the community, it helps to gather all interviewers and supervisors to do the following:

- Last-minute troubleshooting
- Confirm availability of all necessary supplies (see an example of a Field Implementation Checklist on next page)
- Confirm assigned locations of each survey team for that day
- Review community entry protocol (visit local leaders, health workers)
- Review household selection protocol
- Review respondent selection protocol.

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

EXAMPLE OF A FIELD IMPLEMENTATION CHECKLIST
(to be completed by each survey team)

Transportation

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

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

Food and Other Provisions

- ❑ Drinking Water
- ❑ Bag lunch/food allowance
- ❑ First Aid Kit

Survey Equipment

For interviewers

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

For Supervisors

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

The following page contains an example of a Quality-Control Checklist that supervisors can use when observing interviews.

SUPERVISOR INSTRUCTIONS FOR QUALITY CONTROL

Every supervisor should observe **at least one** interview per interviewer **each day**. During the observed interview, try not to interrupt the process. Only fill out the checklist. Once the interviewer has finished the interview, go with the interviewer to another location to discuss any important issues. Remember to highlight his/her strong areas of performance in addition to the areas that might need improvement.

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

KPC QUALITY-CONTROL CHECKLIST

NAME OF INTERVIEWER: _____
NAME OF SUPERVISOR: _____
CLUSTER LOCATION: _____
DATE: ___/___/___
DD/MM/YY

	DID THE INTERVIEWER. . .	PERFORMED CORRECTLY?	
		YES	*NO*
1.	Select the *household* correctly?		
2.	Select the *respondent* correctly?		
3.	Introduce him/herself correctly?		
4.	Read the consent statement at the beginning of the interview?		
5.	Correctly record information on cover page (such as interview date, name of community, mother's/child's name, mother's/child's age/date of birth, child's sex)?		
6.	Speak clearly during the interview?		
7.	Use culturally appropriate body language?		
8.	Have neutral facial expressions/body language (did not react positively or negatively to the respondent's answers)?		
9.	Ask leading questions that might have influenced the respondent's answers?		
10.	Read the questions exactly as they were written?		
11.	Write legibly on the questionnaire?		
12.	Follow the skip patterns correctly?		
13.	Read responses aloud when supposed to?		
14.	Prompt the mother for all answers (say "Anything else?") for questions that allow multiple responses?		
15.	Weigh/measure the child correctly?		
On a scale of 1 (needs follow-up training) to 10 (excellent), I rate the interviewer's performance during this interview as follows (circle one): 1 (needs more training) 2 3 4 5 6 7 8 9 10 (excellent)			

- APPROXIMATE DURATION OF INTERVIEW: _____ minutes
- GENERAL NOTES: _____

_____

SUPERVISOR'S SIGNATURE: _____

KEEPING A RECORD OF FIELD PROBLEMS

Your project can also create a daily reporting form for recording field problems. Below are examples of problems that should be documented as they arise and submitted to the KPC coordinating team daily.

- Problems with household or respondent selection
- Problems with the Questionnaire, for example
  - ⇨ List of questions not understood by mothers
  - ⇨ Incorrect skip patterns
  - ⇨ Inappropriate terminology or wording
- Non-functioning or lost equipment
- Other problems encountered by interviewers

8. TABULATE SURVEY QUESTIONNAIRES

REQUIREMENTS FOR THIS PART OF THE KPC PROCESS

To tabulate questionnaires, you will need the following:

- Completed KPC questionnaires
- A systematic way of checking for errors (Quality-control checklists for field supervisors, computerized error-checking program)
- List of KPC indicators
- Tabulation and analysis plan
- Data entry and analysis programs (for analysis by computer).

Data Cleaning

Before you begin analyzing data, it is first necessary to fix errors that were made during data collection or data entry (if entered into a computer). Data cleaning involves identifying and correcting those mistakes. It is important that both the coordinating team and other field project staff be involved in this activity.

REMINDER

It is not necessary to wait until all of the interviews are completed before you begin cleaning the data. Error checking can take place during field implementation, when it is still possible to correct mistakes. As interviews are completed, supervisors should review the completed questionnaires to make sure that interviewers filled them out correctly. Supervisors can then follow up with interviewers to correct any mistakes identified, revisit respondents (if necessary), and make sure that those same mistakes are not repeated in the remaining interviews.

If there are good quality-control procedures during data collection, then the error-checking process should not be that difficult. As discussed in the previous section on field implementation, it helps to have a special form with a checklist for supervisors

when reviewing questionnaires in the field. The following are examples of mistakes to look for:

- Blank questions
- Wrong codes
- Incorrect skip patterns (going to the wrong question after the respondent gave a particular answer)
- Entering the correct response in an incorrect location
- Unreadable marks.

Once the coordinating team is confident that the data are clean, they can begin to tabulate the data. It is recommended that the tabulation workshop take place one day after data collection has ended. However, tabulation (by hand or computer) can begin in the field while data are still being collected. For example, data can be entered as soon as supervisors turn in reviewed questionnaires to the KPC coordinating team. If your project has opted to do some manual tabulations of the data, this process can also begin in the field. For example, hand-tabulation tables can be organized by cluster, so that survey teams can complete hand-tabulation tables as they complete interviews in each cluster.

There should be quality control whenever there is the *potential* to make errors. If the project decides to conduct the analysis by computer, it is important to have a system of checking for errors during the data entry process. Persons responsible for entering data will need to be closely monitored. It helps to divide data entry personnel into teams of two rather than have them work separately. For each completed questionnaire, one person can read the respondent's answers, while the other person enters the correct codes into the computer. It also helps to have a double entry system of entering data. With double entry, the data are first entered by one person, then entered a second time by a different person. Any discrepancies can therefore be resolved before running the computerized analysis program.

Why Do We Need a Tabulation Workshop?

Once the proper files are created and the data are entered, computer analysis can be fairly quick and will allow projects great flexibility in terms of the type of analyses that can be conducted. However, people often have difficulties understanding how information collected from individual respondents translates into numbers and percentages generated by a computer. For that reason, it is a good idea to have a Hand Tabulation Workshop with persons who were either involved in the survey or have a stake in the survey's results (and the project in general). This is recommended even if your project has decided to analyze the KPC by computer. Manual (hand) tabulation gives a hands-on feeling for what the data mean to a larger number of people. If local partners and stakeholders opted not to be directly involved in data collection activities, invite them again to participate in the Hand Tabulation Workshop. This is a prime opportunity for all stakeholders to work directly with the data and identify and prioritize problems as a group. By being transparent in terms of how the data are collected, analyzed, and interpreted, a project can use the Hand Tabulation Workshop to build consensus among stakeholders. Hand tabulation is also a good way to validate the results generated by a computer.

At the workshop, encourage participants to think aloud in terms of their reactions to, and possible explanations for, the findings.

Who Should Attend the Tabulation Workshop?

All KPC coordinating team members should be present at the KPC Tabulation Workshop and can facilitate the hand tabulation process. However, other individuals should have the opportunity to work directly with the data. As a courtesy, invite other individuals who either have a stake in the project or are affiliated with agencies and institutions that work in the same geographic area. Persons who were involved in the survey, namely supervisors and interviewers, can be invited as tabulators. You do not have to limit participants to individuals who work in the health field only. For example, local communication specialists, water and sanitation experts, or qualitative researchers may also be invited.

In planning the Hand Tabulation Workshop, keep in mind that a group that is too large will be hard to facilitate. Aim to have a group that is manageable—given time, space, and other constraints—yet includes people who represent different perspectives on child health and survival. With a diverse group of workshop participants, your project might gain insight into why certain problems exist and how those problems can be addressed.

REMINDER

In the informed consent statement read at the beginning of each interview, the interviewer assured the respondent that the information she provided would be kept confidential. It is highly important that steps be taken during the KPC Tabulation Workshop to maintain the confidentiality of all respondents and their children. Personal identifiers should not be disclosed to workshop participants who were not directly involved in the survey.

It is suggested that at the beginning of the workshop, the facilitators remind participants that they must respect the each respondent's right to privacy, and that they should not discuss information regarding any particular respondent with others outside of the workshop.

What Should Be Covered in the Tabulation Workshop?

It is not necessary to tabulate every single question from the survey. In addition, if the coordinating team has decided to conduct the analysis by computer, it is not necessary to hand tabulate all of the KPC indicators during the workshop. The objective of the Hand Tabulation Workshop is to ensure that other individuals understand the data, see value in it, and ultimately use it to prioritize problems and develop solutions.

Things to Consider

- If time permits, choose at least one indicator for each of the project's technical interventions.
- To identify indicators, refer to your KPC analysis plan, which should have been developed during the pre-implementation stages of the survey.
- Prepare hand tabulation tables for each of the key indicators, in case workshop participants express an interest in tabulating particular indicators.

REMINDER:

Each *KPC2000+* module contains at least one example of a hand tabulation table.

- After the group has hand tabulated a few indicators, it is helpful to explore subgroup differences. For example, are certain indicators different for boys than for girls? For children of young women than those of older women? For children whose birth was inadequately spaced versus those whose birth was adequately spaced? *These and other extended analyses will be discussed in the next section of the field guide* ("Analyze the Data").

Reminders When Planning a Hand Tabulation Workshop

- Identify which indicators will be hand tabulated at the workshop.
- Reach a consensus on how those indicators are defined before the workshop.
- Determine which survey questions pertain to each indicator.
- Create and photocopy tabulation tables for workshop participants.
- Devise clear-cut instructions for workshop participants on how to tabulate indicators by hand.
- Give invitees ample notice in terms of where and when the workshop will be held.
- Reserve an adequate facility where the workshop can be held (one that has a reliable power source, if one is needed, and enough table space to organize and review survey forms and record responses on a tabulation table).
- Obtain all necessary materials such as markers/chalk, pens, large sheets of white paper/chalkboard.
- Have all completed questionnaires present at the workshop.
- Identify at least one person on the coordinating team who will be responsible for

documenting the process (e.g., take notes, videotape the workshop).

REMINDER

If your project will be tabulating all indicators by hand, the coordinating team should allot a sufficient amount of time, personnel, and other resources to accomplish this accurately and efficiently.

During the workshop, members of the coordinating team should review each team's output and provide feedback. After hand tabulating indicators in small groups, everyone can reconvene as one large group to share results based upon their tabulations. If only a few indicators were hand tabulated, the coordinating team can explain to the group that in the interest of time, only a few indicators will be hand tabulated. At that time, they can share results for the other key indicators that were generated by computer (if available). Once that process is complete, the group can begin analyzing and discussing findings from the survey.

9. ANALYZE THE DATA

REMINDER

It is important to identify errors that occurred in the field or during data entry before you start analyzing the data.

The current version of the Analysis section presents advice on how to analyze data that are aggregated for an entire program area. However, if your project chooses to select samples from different program units (e.g., lots), it is important to explore differences between those program units. The next version of the field guide will include discussions on how to analyze data from KPC surveys that used LQAS or a methodology other than cluster sampling.

The first step in the analysis process is to compute the survey indicators according to their standard international definitions. If the country where you are working has defined some of those indicators differently, it is suggested that you do a second set of calculations based on those national definitions. The first way ensures that your indicators comply with international standards. However, the second way is important if your project wants to compare its KPC findings with other data in the country.

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

This section of the field guide suggests ways to analyze the data so that a project can develop a more targeted response to the issues and problems faced by its target population. In this section, the Analysis Phase has been divided into the following:

- Identifying Problems
- Analyzing Problems.

Before discussing each of these tasks, the following page presents a list of issues to consider if your project has decided to analyze the data using a computer.

THINGS TO CONSIDER WHEN ANALYZING DATA BY COMPUTER

1. Number and qualifications of persons who will be involved in data entry.

2. Number of computers needed.

3. Reliability of power sources for computers, particularly in areas where there are frequent power outages.

4. A system of backing up computer files.

5. Software requirements: Epi Info is public-domain software that is free of cost. The software and its manual can be downloaded from the Centers for Disease Control Web site (www.cdc.gov/epiinfo). Remember that it takes a lot of time to create the necessary files for entering, checking, and analyzing the data. Therefore, it is important to consider these tasks when planning and scheduling post-implementation activities. Once there is a final draft of the questionnaire, staff can start working on the computer files. Staff can enter the data as completed questionnaires are turned in by the field supervisors. By the end of data collection, all necessary computer files and programs should be created and ready for use. It is important that the computerized questionnaire file resembles the hard-copy questionnaire as much as possible.

6. Analysis plan: To write an analysis program, it is necessary to know what types of indicators will be generated during the analysis. The KPC coordinating team should have identified the KPC indicators when developing the survey. It is important to have the analysis plan in writing, which will serve as an easy reference during the Analysis Phase as well as in the future.

7. Keep in mind that software packages such as Epi Info are usually available in English, Spanish, and French. If your KPC questionnaire had to be translated into a local language or dialect, the persons involved in data entry and analysis will need to be versed in one of those three major languages, in addition to the language in which the questionnaire was translated.

8. Instant access to data at the Analysis Workshop: If possible, setup one or two computers with the analysis software and a computerized file containing the KPC data. In doing so, there is the potential to perform additional cross tabulations suggested by workshop participants immediately upon their request.

IDENTIFYING PROBLEMS

Before examining the study indicators, examine the frequencies of responses to each survey question. What is the general profile of the sample? For example, what is the age distribution of respondents? Do most of them earn wages outside the home? What is their level of education? How many households use piped drinking water? What type of sanitation facility do most households have?

Next, compute the *Rapid CATCH* indicators and any other indicators decided on by your project.

- Examine the key findings from the survey. Do they confirm what the project and stakeholders expected? Which results are surprising or troubling? **In the Analysis Workshop, encourage participants to think aloud and share their impressions with the rest of the group.**

- Compare key results with data from available standards. Standards can come from project objectives, national objectives, national or international health messages, or reported data from other sources (national or local). Does the KPC paint the same picture as data from other sources? If not, what are some possible explanations? Based on the KPC, how does the program area compare with neighboring areas or the country as a whole?
- Identify gaps between current practices and coverage.

ANALYZING PROBLEMS

Exploring whether differentials exist for certain indicators is an important step in problem analysis. Once the key indicators are computed, workshop participants can begin to cross tabulate the data according to select background variables. Even if your project conducts the analysis by computer, it helps to draw 2x2 tables on large sheets of paper so that workshop participants can get hands-on experience in data exploration.

For every survey finding, ask workshop participants to consider the following question: Does this finding apply to the population as a whole, or might there be important differences between certain population groups? KPC sample sizes are typically small, but results from the cross tabulations might highlight differentials that warrant further investigation.

The following table presents some background characteristics to consider when analyzing the data, as well as examples of ways to categorize mothers and children. It is not necessary to explore differences by each of the characteristics listed in the table. In addition, there might be other variables not listed in the table that are relevant to your project. Do what makes sense for your local context.

REMINDER

If your project is interested in drawing comparisons between certain populations groups, this should be reflected in the questionnaire and/or the sampling strategy.

EXAMPLES OF CHARACTERISTICS TO CONSIDER WHEN ANALYZING KPC DATA

	CHARACTERISTIC	SUBGROUP CATEGORIES	COMMENTS
1.	Maternal age	>25 years, ≥ 25 years	If you are working in an area where there is a lot of premarital childbearing or where women marry at a young age, you might want to use a younger age cut-off (for example, compare mothers who are < 20 years with those who are ≥ 20 years.)
2.	Child's age	0–11 months, 12–23 months	
3.	Sex of child	male, female	
4.	Under-5 household density	Households with fewer than two children under 5, households with two or more children under 5.	Even if all of the children in the household do not belong to the same mother, the number of young children within the same household gives an indication of whether there might be competition for resources such as food.
5.	Previous birth interval	Children who were born less than 24 months after the previous surviving sibling, children who were born 24 or more months after their previous surviving sibling.	Because the KPC2000+ is based solely on children who are alive at the time of the interview, information collected on birth intervals may not exactly match information from other data sources (in particular, those that take non-live births and/or both surviving and non-surviving children into account). Nevertheless, the length of the previous birth interval might be important to consider when examining infant/child feeding practices, nutritional status, and other factors.
6.	Type of caregiver	Maternal, non-maternal (such as the grandmother or an older sibling).	This information is not in the standard KPC modules. However, if your target population has a large number of orphans (e.g., due to war or the HIV/AIDS epidemic), or it is common for biological mothers to leave their children with other individuals, this might be an important variable to consider.
Your project might include additional questions in the survey to explore differences in health outcomes by other variables (such as mother's marital status (married versus unmarried), type of household (e.g., female-headed versus male-headed), or socioeconomic status (e.g., household owns television and radio versus household that does not)). Your project should include whatever context-specific information it feels is relevant to the way it plans and manages its intervention activities.			

To make comparisons, you need the appropriate data. This is why it is important to develop a preliminary analysis plan before developing the KPC questionnaire or designing the sampling strategy. If you postpone developing an analysis plan, you might find that the information you need was not included in the survey. For example, if you would like to compare the outcomes of children whose mothers are currently married versus unmarried, you would need to do one of the following:

- Include a question on marital status in your questionnaire (Although, there is a chance that you might not have enough cases in each group to make meaningful comparisons)
- Make provisions in your sampling strategy to select both married and unmarried mothers for the study.

Because the number of cases in a KPC survey is generally small, do not rely on results from statistical tests (p-values, chi-square statistics) to flag differentials that warrant further investigation or attention.

A small sample size limits how in-depth the analysis can be. Even if you are looking at indicators that are based on the entire sample, try to avoid comparing more than two categories. For example, there may not be enough cases to explore differences between women in 5-year groups (15–19, 20–24, 25–29, 30–34, etc.). As a result, collapse the 5-year age groups into two broad age categories (such as <25 and ≥ 25).

Below is a 2x2 table that illustrates how to divide the cases in your sample according to the result you are measuring (KPC indicator) and another variable of interest.

KPC Result (For example, the percentage of children with pneumonia who were taken to a health facility)

CHARACTERISTIC OF INTEREST (For example, is the child's primary caregiver his/her biological mother?)		**YES**	**NO**	**TOTAL**
	YES	Number/percent of children with the characteristic and the desired result (Mother is the primary caregiver, child was taken to a health facility for his/her pneumonia)	Number/percent of children with the characteristic, but not the desired result (Mother is the primary caregiver, child was not taken to a health facility for his/her pneumonia)	Total number of children with the characteristic
	NO	Number/percent of children who do not have the characteristic, but have the desired result (Mother is not the primary caregiver, child was taken to a health facility for his/her pneumonia)	Number/percent of children who do not have the characteristic, or the desired result (Mother is not the primary caregiver, child was not taken to a health facility for his/her pneumonia)	Total number of children without the characteristic
	TOTAL	Total number of children with the desired result	Total number of children without the desired result	Total Number of Children with Pneumonia

As you explore the data, additional questions may arise. Can answers to those questions be found in the KPC? Would qualitative research explain the nature of the problem?

Qualitative research can yield important information on community knowledge, beliefs, and normative practices that affect child health and survival. When used after KPC survey data have been collected and analyzed, qualitative research might provide explanations for phenomena that were identified but not sufficiently explained by the KPC.

10. WRITE THE SURVEY REPORT

This updated version of "Writing the Survey Report" (initially written by the PVO CSSP) provides recommendations on the format and content of the KPC survey report.

Why Do We Need a Survey Report?

The KPC survey report is an important outcome of the KPC process. It should provide a detailed description of the study, present survey findings, and discuss the program implications of those findings. Individuals who were not involved in the study should be able to read the report and get a good sense of the process and methods, not just the major findings.

Projects are encouraged to share the report with partner organizations, donor agencies, and other agencies/institutions working in the same geographic area. If the KPC was part of the project's baseline assessment, then the survey report can serve as background material during the DIP-writing workshop. In addition, project staff can work with local partners and stakeholders to make the survey report a living document. For example, they can experiment with ways to publicly display KPC findings and can explore innovative ways to track result-level information during the life of the project. This will be discussed in greater detail in the section of the guide entitled "Use KPC Data".

When Should We Prepare the Survey Report?

The coordinating team should aim to have a first draft of the survey report by the end of the analysis workshop (i.e., about 2 days after data collection has ended). The team can then devote a few days to revisions, producing the final report within one week of completing data collection. This is a realistic expectation, especially if your team starts drafting sections of the report on the process, tools, and methodology during the pre-implementation phase of the survey. Also, if local partners and stakeholders are actively involved in the planning, conduct, and analysis of the survey, then their comments should be reflected in the first draft of the report. It will therefore be less likely that they will have many additional comments that need to be incorporated in the final draft.

REMINDER

In addition to developing local capacity to conduct rapid surveys, it is equally important to develop skills in documenting and disseminating information. If your project has hired a consultant to act as Survey Coordinator, the consultant should work closely with members of the KPC coordinating team when drafting the survey report.

I. FORMAT

It is useful to present material in the following manner:

- Cover Page (include title, date, PVO/country, partner organizations, other participating organizations, and author names)
- Acknowledgments (list all supervisors, interviewers, etc., and their titles)
- Table of Contents
- Executive Summary (written last)
- Background
- Process and Partnership Building
- Methods
- Results
- Discussion
- Bibliography
- Appendices.

II. CONTENT

Background

This section of the report should include background information on the context in which the PVO is working. Examples of relevant information are as follows:

- Project location
- Characteristics of the target beneficiary population

- Health, social, and economic conditions within the project area
- National standards/policies regarding maternal and child health (e.g., immunization, breastfeeding, or Safe Motherhood policies).

Much of the above information is probably in your project proposal or DIP. However, you can also include relevant information from other sources identified during the Pre-Implementation Phase (refer to the section entitled "2. Determine Needs and Information Gaps" in the Pre-Implementation section of this guide).

The authors should also give an overview of the project, namely the following:

- Goals
- Objectives
- Intervention activities
- Objectives of the KPC survey.

Process and Partnership Building

As a reminder, the KPC was designed to be participatory, engaging local partners and stakeholders in information gathering, analysis, and decisionmaking. Participatory research is conducive to partnership and capacity building. In addition, it fosters a sense of local ownership of KPC results and greater use of information for local decisionmaking.

This version of the KPC report-writing guidelines stresses the importance of discussing partnership and capacity-building activities as they relate to the KPC survey. A discussion of the following would be useful:

- Methods of identifying and engaging local partners/stakeholders in the KPC
- Specific roles of local partners/stakeholders in the KPC survey
- Constraints in making the KPC process more participatory
- Innovations in partnership building and participatory research used in the study.

Partnership building is a dynamic and sometimes intensive process. If your project encountered certain barriers to making the process participatory, or if the project used an innovative approach to maximize local involvement, it will be interesting to document this in the survey report.

Methods

In the Methods section of the survey report, it is important to discuss the following:

- Questionnaire
- KPC indicators
- Sampling design
- Training
- Data collection and quality control procedures
- Data management/data analysis.

Questionnaire:

- Questionnaire development process
- Scope of the survey (topic areas covered)
- Survey length
- Versions of the questionnaire (if more than one type of respondent is sampled)
- Specially adapted survey questions
- Translation into local languages

Study indicators:

- List of indicators grouped by intervention or topic area
- Definition of each indicator

In adapting the questionnaire, remember that the KPC should yield a concise and manageable set of indicators for reporting results. Projects will probably need to consult additional data sources and employ other information-gathering techniques to obtain all the information necessary for developing project activities or strategies. As a reminder, the *Rapid CATCH* yields indicators that are useful to all CS projects, regardless of their intervention mix. Reporting on these indicators is strongly encouraged by the CORE Monitoring and Evaluation Working Group.

Sampling design:

- Sources of population data (used in the selection of sample areas)

- Type of design used (e.g., 30-cluster sampling or Lot Quality Assurance Sampling designs; parallel sampling techniques)
- Sample size calculations
- Selection process

As stated earlier in this guide, 30-cluster sampling has been traditionally used with KPC surveys, but it is not the only method of sampling. The survey report should discuss details of the sampling process, namely a) the type of design used; b) the process used to select the sample areas (clusters/lots), households, and respondents; c) the number of sample areas; and d) the number of interviews conducted within each sample area. It is also helpful to state any selection protocols employed during the study (e.g., what interviewers were instructed to do when there was more than one child less than 2 within the same household).

It is also helpful to include information on the number (and when available, key characteristics) of women who refused to be interviewed.

Training:

- Selection of interviewers—process of selection and general profile of supervisors and interviewers (e.g., female, high-school educated, staff from partner organizations)
- Training of supervisors and interviewers (duration of training, person(s) who conducted the training, content/structure of training sessions)
- Strengthening local capacity to conduct future small-sample surveys

Data Collection:

- Average length of interview
- Number of days for data collection
- Major constraints/field problems
- Quality-control procedures

Sometimes there are unforeseen circumstances that impact the progress of fieldwork. Describe major problems encountered during the fieldwork and discuss the potential impact of those problems on data quality. Also discuss the steps that were taken to maintain high data quality in the field.

Data Analysis

- Method of data analysis (i.e., hand tabulation and/or computer tabulation)
- Statistical software packages used, if any
- Description of person(s) involved in data management/analysis (e.g., supervisors/

interviewers, PVO field staff, MOH personnel)

- Quality-control procedures (e.g., error checking during the data entry process)
- Hand-tabulation workshops, if any

Results

This section of the report should present the results for each of the study indicators. It is helpful to both readers and report writers to present findings in the form of tables and to refer to these tables within the text. For the first draft of the report, which is usually written in the field immediately following the survey, it is acceptable to include the frequency distributions for each of the survey's questions. In the final version of the report, it is not necessary to include frequencies for each survey question. They may, however, be included in the appendix of the report. Cross tabulation of the data by key variables (e.g., sex of child, maternal age) are encouraged. Although KPC sample sizes are typically small, cross tabulations might suggest important differences between subgroups of mothers/children that warrant further investigation. It is not necessary to present a table for each cross tabulation. However, it is helpful to report findings for any cross tabulations that are performed, even if it can only be stated that no differences were observed for certain variables. The following is an illustration of how to present cross-tabulated data.

Indicator: percentage of mothers of children 0–23 months who know at least two signs of childhood illness that indicate the need for treatment, by maternal age group

MATERNAL KNOWLEDGE OF AT LEAST TWO SIGNS OF CHILDHOOD ILLNESS INDICATING THE NEED FOR TREATMENT

		YES	NO	TOTAL	PERCENT
MATERNAL AGE	<25 years				
	≥25 years				
	Total				

Readers of the report should clearly understand the numerator and denominator of each indicator. It is very helpful to have a single table at the beginning of the Results section that lists all indicators, their numerators, denominators, percents, and confidence limits. The table on the following page is an illustration.

EXAMPLE OF A SUMMARY TABLE OF SURVEY INDICATORS

INDICATOR	NUMERATOR	DENOMINATOR	PERCENT	CONFIDENCE LIMITS
Percentage of children 0–23 months who are underweight (-2 SD from the median weight-for-age, according to the WHO/NCHS reference population)				
Percentage of children 0–23 months who were born at least 24 months after the previous surviving child				
Percentage of children 0–23 months whose births were attended by skilled health personnel				
Percentage of mothers with children 0–23 months who received at least two tetanus toxoid injections before the birth of their youngest child				
Percentage of children 0–5 months who were exclusively breastfed during the last 24 hours				
Percentage of children 6–9 months who received breastmilk and complementary foods during the last 24 hours				
Percentage of children 12–23 months who are fully vaccinated (against the five vaccine-preventable diseases) before the first birthday				
Percentage of children 12–23 months who received a measles vaccine				
Percentage of children 0–23 months who slept under an insecticide-treated net (in malaria risk areas) the previous night				
Percentage of mothers with children 0–23 months who cite at least two known ways of reducing the risk of HIV infection				
Percentage of mothers with children 0–23 months who report that they wash their hands with soap/ash before food preparation, before feeding children, after defection, and after attending to a child who has defecated				
Percentage of mothers of children 0–23 months who know at least two signs of childhood illness that indicate the need for treatment				
Percentage of sick children 0–23 months who received increased fluids and continued feeding during an illness in the past two weeks				

A WORD ABOUT CONFIDENCE LIMITS

The KPC results serve as best estimates of key child survival indicators. It is important to remember that estimates from any survey are associated with a certain level of error. For each indicator, the study estimate is just one value within a range of possible values. When reporting KPC results, it is helpful to state the *confidence limits*, which indicate the margin of error for each survey finding. Project staff can use confidence limits to better compare survey findings with the project's objectives, reported national levels, or findings of other similar surveys. Below are two examples of how to use and interpret confidence limits.

Using confidence limits—Example #1

- Objective: By the end of the project, 80 percent of infants will be fully immunized in their first year of life.
- Indicator: Percentage of children 12–23 months who received the required number of immunizations (according to national/international protocol) against vaccine preventable diseases before the first birthday.
- KPC result: 65 percent of children 12–23 months in the survey were fully immunized by age 1 year. The confidence limits are calculated as plus or minus 10 percent.
- Conclusion: We are 95 percent confident that the true proportion of fully immunized children in the population is between 55 percent and 75 percent (65% ±10%). The best estimate of the true proportion is 65 percent.
- Discussion: A comparison of the survey finding—including its confidence limits (the margin of error)—with the project objective indicates that the project did not achieve its objective. That is, the evidence suggests that the true proportion of fully immunized children in the population is less than 80 percent (the probability that 80 percent is the true proportion in the population, given the survey finding, is less than 5 percent). The project should study other EPI findings from the survey to identify the barriers to achieving the objective. The project should also consider using qualitative research methods to shed further light on the matter. For example, low coverage of a specific antigen (e.g., measles vaccine) may be the principle barrier to full immunization. Or, there may be a high drop out rate between the first DPT (or OPV) and the third.

Using confidence limits—Example #2

- Objective: By the end of the project, 80 percent of infants will be fully immunized in their first year of life.
- Indicator: Percentage of children 12–23 months who received the required number of immunizations (according to national/international protocol) against vaccine preventable diseases before the first birthday
- KPC result: 76 percent of children 12–23 months in the survey are fully immunized. The confidence limits are calculated as plus or minus 9 percent.
- Conclusion: We are 95 percent confident that the true proportion of fully immunized children in the population is between 67 percent and 85 percent (76% ± 9%). The best estimate of the true proportion is 76 percent.
- Discussion: A comparison of the survey finding—including its confidence limits (the margin of error)—with the project objective indicates that the survey finding is consistent with the objective. However, the best estimate of the true proportion is lower than the objective. It is likely, therefore, that the project did not completely achieve its objective of 80 percent coverage. The project should study other EPI findings to determine specific areas for improvement (low coverage of specific antigens like measles vaccine; dropout rate for DPT or OPV is higher than expected) to further increase immunization coverage. Qualitative research might also shed light on the "hows" and "whys" behind lower-than-expected EPI coverage.

Confidence Limit Formulas

Confidence limits with a SRS

The formula for calculating the confidence limits of a survey finding when using SRS is:

$$\mathbf{P = p \pm Z\delta \text{ where } } \delta = \sqrt{(pq/n)}$$

Z = 95 percent confidence = 1.96
P = true proportion in the population
p = proportion found in the survey
q = 1-p
n = size of sample or sub-sample

EXAMPLE: Assume **p** = .4, **q** = .6, **n** = 210, **z** = 1.96

$$\mathbf{P} = \mathbf{p} \pm Z \times \sqrt{(pq/n)}$$

$$\mathbf{P} = \mathbf{p} \pm .07$$

$$\mathbf{P} = .4 \pm .07 = .33 \leq \mathbf{p} \leq .47$$

Conclusion: We are 95 percent confident that the true proportion in the population is between 33 percent and 47 percent. The best estimate for the true proportion in the population is 40 percent.

Table 1. Confidence Limits for a SRS:

$$P = p \pm z\sqrt{(pq/n)}$$

p	Sample Size (n)				
	180	210	240	270	300
0.05	±.03	±.03	±.03	±.03	±.02
.2	±.06	±.05	±.05	±.05	±.05
.4	±.07	±.07	±.06	±.06	±.06
.6	±.07	±.07	±.06	±.06	±.06
.8	±.06	±.05	±.05	±.05	±.05
.95	±.03	±.03	±.03	±.03	±.03

Confidence limits with a cluster sample

Cluster sampling methods often provide survey findings that are less precise than the findings obtained using SRS. This comes from the potential bias of sampling in groups (of households or individuals) rather than sampling individuals. Sampling in groups presents a possible bias because behavior among group members is more likely to be similar. A sample of these groups, therefore, may not be as representative of the entire population under study as a sample of randomly selected individuals. The implication of this bias is that the confidence limits of a finding from a cluster survey are often wider than the confidence limits of a finding from a SRS, all other things being equal.

Calculating cluster survey confidence limits by computer: Computer software programs such as Epi Info can easily calculate the confidence limits for a finding from a cluster survey. **Note that computerized survey forms need to have a field identifying the respondents cluster (cluster id) to calculate confidence limits by computer.**

Calculating cluster survey confidence limits by hand: The formula for calculating the confidence limit of a cluster survey finding by hand is:

$$P = p \pm z\sqrt{(pq/n')}$$, where n' = the ***effective sample size*** of the sample or sample

Effective Sample Size (n') = n/e, where:

- **n** = size of survey sample or sub-sample
- **e** = ***design effect***. The design effect is a value corresponding to how much the cluster survey departs from the assumptions of a SRS. The design effect is used to correct the value of **n** used to calculate the confidence limit of a cluster survey.

Design Effect—To calculate the confidence limit by hand, projects usually estimate the value of the design effect. This is because the formula for calculating the design effect is difficult to do by hand and is most often done by computer. For the variables in the KPC survey, the design effect usually ranges in value between 1 and 2. Projects can estimate the confidence limits of a finding with the following methods:

- Calculate the confidence limit of a finding assuming the design effect is 1 (no difference in precision between cluster sampling and a simple random sample). Then, calculate the confidence limit again, this time assuming that the design effect is 2 (the cluster survey sample size needs to be twice as large to maintain the precision of a simple random sample). Finally, report both confidence limits as the range of possible values.
- Calculate the confidence limit of a finding assuming the design effect is 2 (the cluster survey sample size needs to be twice as large to maintain the precision of a simple random sample). This is a conservative estimate as the true design effect will often be less than 2.
- If confidence limits for the same or similar finding are available from other local cluster surveys (e.g., EPI Cluster Survey), use the design effect reported for that survey to calculate the confidence limit. Report the source of data for the design effect value used in the survey report.

EXAMPLE: Assume **p** = .4, **q** = .6, **n** = 210, **design effect (e)** = 2, **z** = 1.96

$$\mathbf{P} = \mathbf{p} \pm Z \times \sqrt{(pq/n')}$$
$$\mathbf{P} = \mathbf{p} \pm 1.96 \times \sqrt{[(.6 * .4)/(210/2)]}$$
$$\mathbf{P} = \mathbf{p} \pm 1.96 \times \sqrt{(.24/105)}$$
$$\mathbf{P} = \mathbf{p} \pm .09$$
$$\mathbf{P} = .4 \pm .09 = .31 \leq \mathbf{p} \leq .49$$

Conclusion: We are 95 percent confident that the true proportion in the population is between 31 percent and 49 percent. The best estimate for the true proportion in the population is 40 percent.

Table 2. Confidence Limits for a Cluster Survey: Assume Design Effect = 1.5

n (n′ = n/1.5)

p	180 (120)	210 (140)	240 (160)	270 (180)	300 (200)
.05	±.04	±.04	±.03	±.03	±.03
.2	±.07	±.07	±.06	±.06	±.06
.4	±.09	±.08	±.08	±.07	±.07
.6	±.09	±.08	±.08	±.07	±.07
.8	±.07	±.07	±.06	±.06	±.06
.95	±.04	±.04	±.03	±.03	±.03

Table 3. Confidence Limits for a Cluster Survey: Assume Design Effect = 2

n (n′ = n/2)

P	n (n′)				
	180 (90)	210 (105)	240 (120)	270 (135)	300 (150)
.05	± 0.05	± .04	± .04	± .04	± .03
.2	± .08	± .08	± .07	± .07	± .06
.4	± .10	± .09	± .09	± .08	± .08
.6	± .10	± .09	± .09	± .08	± .08
.8	± .08	± .08	± .07	± .07	± .06
.95	± .05	± .04	± .04	± .04	± .03

Calculating Confidence Limits for an LQAS KPC Survey

LQAS yields a small number of cases within each program management area (lot). As a result, lot-specific coverage estimates and confidence limits will not be precise, and will therefore provide information that is meaningless. However, it is possible to calculate a coverage estimate (and corresponding confidence limits) for the project area as a whole with a great deal of precision by combining lots. In doing so, it is important to remember that the number of beneficiaries probably varies from one lot to the next. In these instances, when

calculating a program-wide coverage estimate, you should consider weighting results from each lot by the total number of beneficiaries residing in that lot.

*******DISCLAIMER: Although weighted estimates are regarded as more accurate than unweighted estimates, in reality, the difference between weighted and unweighted estimates is usually not that large.*******

How to calculate a weight for each lot:

Definition of Symbols:

n = LQAS sample size (The total number of children in your sample)

n_i = sample size for a particular lot (**19** children in each lot is often used)

N = total number of children in the project area

N_i = total number of children in a particular lot

The weight for a given lot (w_i) = Total number of children in that lot divided by total number of children in the project area

= N_i / N

In other words, the weight is simply the proportion of the program area's total population that lives in a particular lot.

Table 4 provides an example of calculating weights. In this example, assume that the project has five supervision areas that it has designated as lots.

Table 4. Determining Weights for LQAS Lots

LOT (SUPERVISION AREA)	POPULATION SIZE	WEIGHT (w_i)
Supervision Area A	1,600	1,600/10,000 = **.16**
Supervision Area B	2,300	2,300/10,000 = **.23**
Supervision Area C	2,200	2,200/10,000 = **.22**
Supervision Area D	2,000	2,000/10,000 = **.20**
Supervision Area E	1,900	1,900/10,000 = **.19**
TOTAL (Entire Project Area)	10,000	

You can now use these weights to calculate a coverage proportion for your entire project area.

How to calculate a coverage proportion:

- For each lot sample, divide the number of children who have the characteristic of interest by the total sample size of that lot.

- Multiply the number calculated in step 1 by the weight that you calculated for that lot (as done in Table 4). Do this for each lot.
- Sum the numbers that were calculated in Step 2 across all lots. The final number is the coverage proportion for your entire sample area.

Below is an example of how to calculate a coverage proportion. For this example, suppose that the project wants to assess full immunization coverage for the entire project area. Assume that the same lots and weights calculated in the previous section (under "How to calculate a weight for each lot") apply to this example.

Definitions of Symbols:

n_i = LQAS sample size (The total number of children in your sample)

x_i = Number of children within lot *i*'s sample who are fully immunized before the first birthday

w_i = Weight for lot *i* (see previous section "How to calculate a weight for each lot")

Table 5. Weighting Data from Each Lot to Determine Overall Coverage

LOT	n_i	x_i	x_i/n_i	w_i	$w_i * (x_i/n_i)$
Supervision Area A	19	3	.16	.16	.03
Supervision Area B	19	5	.26	.23	.06
Supervision Area C	19	5	.26	.22	.06
Supervision Area D	19	7	.37	.20	.07
Supervision Area E	19	11	.58	.19	.11
TOTAL	**95**				**.27**

As seen above, the full immunization coverage for the entire project area is 27 percent (.27 x 100).

You can then calculate a confidence limit for the overall coverage estimate using the following formula.

Z = 95 percent confidence = 1.96
P = true proportion in the population
p = coverage proportion for entire program area (see Table 5)
$\mathbf{p_i}$ = coverage proportion for a particular lot (= x_i/n_i in Table 5)
$\mathbf{q_i}$ = $1-p_i$
$\mathbf{n_i}$ = size of sample in a particular lot

$$P = p \pm Z \times \sqrt{\sum \frac{w_i^2 \times (pq/n)}{n_i}}$$

Discussion

In the Discussion, authors are encouraged to do the following:

- Relate key findings from the KPC survey to data from other sources
- Discuss the programmatic implications of the survey findings
- Identify next steps in information gathering
- Present an action plan for community feedback and dissemination of findings.

1. External comparisons

Examples of useful data sources to compare with the KPC are the following:

- Demographic and Health Survey data for the country in which you are working
- Other local surveys
- MOH statistics
- MOH objectives or standards
- PVO's own project objectives
- Reported national data
- WHO/UNICEF objectives or standards.

2. Programmatic Implications

After presenting results and comparing survey data to other comparable data, discuss the implications for the project. Recommendations can be included.

3. Additional Information Gathering

After the KPC data have been analyzed, it might be necessary to conduct qualitative research

to better understand some of the issues raised by the KPC survey. To illustrate the importance of conducting a qualitative follow up, suppose your KPC survey indicates that measles vaccination coverage is very low. Through qualitative research, you might find out that there is a cultural/religious belief that diseases that are contracted only once during a person's lifetime are of spiritual significance, and attempts should not be made to prevent those diseases. This information would be invaluable when considering ways to improve the level of measles vaccine coverage in the target population.

4. Information Dissemination

Describe planned and/or completed activities for feedback of the survey results to communities, the MOH, USAID, and other relevant parties. Immediate feedback following the survey ensures that survey findings are shared at a time when there is peak interest in the findings. It is unlikely that interest in the survey findings will be high if the PVO waits several weeks or months after the survey to provide feedback. If the first draft of the survey report is completed within several days following the survey, then it will be available for handing out at feedback sessions that take place immediately following the survey.

Projects are encouraged to make arrangements for feedback during initial preparations for the survey. If arrangements are not made well in advance, it will be difficult to bring together persons who want to discuss the survey findings immediately after the survey. Once interest in the survey lowers, it will be even more difficult to bring groups together for feedback meetings.

Bibliography

This section will help readers to repeat the methodology for other surveys. Include in this section the source of population data and other sources drawn from for the survey methodology and for comparison data used in the Discussion section of the report. Other useful references (survey research texts, journal articles, manuals, or other publications) can also be listed.

Appendices

The following list of appendices will help readers of the report to answer additional questions that they may have after reading the formal report:

- Manual tabulation tables or computer tables (raw data) for each question
- Population data used to select 30-clusters (or other units of selection if cluster sampling was not used)
- Survey questionnaire in both English and the local language
- A list of persons involved in the survey and their roles

- Training schedule for supervisors and interviewers
- Project resource requirements of the survey, namely monetary costs and amount of staff time devoted to KPC activities.

CONFIDENCE LIMITS FOR SMALL SAMPLE SIZES, ASSUMING A DESIGN EFFECT OF 1.5 OR 2.0

Design Effect=1.5												
		N'=N/1.5										
		N	**30**		**60**		**90**		**120**		**150**	
			(if N=30, N'=20)		(if N=60, N'=40)		(if N=90, N'=60)		(if N=120, N'=80)		(if N=150, N'=100)	
		N'	**20**		**40**		**60**		**80**		**100**	
	0.05		0.10		0.07		0.06		0.05		0.04	
	0.2		0.18		0.12		0.10		0.09		0.08	
P	0.4		0.21		0.15		0.12		0.11		0.10	
	0.6		0.21		0.15		0.12		0.11		0.10	
	0.8		0.18		0.12		0.10		0.09		0.08	
	0.95		0.10		0.07		0.06		0.05		0.04	
Design Effect=2.0												
		N'=N/2.0										
		N	**30**		**60**		**90**		**120**		**150**	
			(if N=30, N'=15)		(if N=60, N'=30)		(if N=90, N'=45)		(if N=120, N'=60)		(if N=150, N'=75)	
		N'	**15**		**30**		**45**		**60**		**75**	
	0.05		0.11		0.08		0.06		0.06		0.05	
	0.2		0.20		0.14		0.12		0.10		0.09	
P	0.4		0.25		0.18		0.14		0.12		0.11	
	0.6		0.25		0.18		0.14		0.12		0.11	
	0.8		0.20		0.14		0.12		0.10		0.09	
	0.95		0.11		0.08		0.06		0.06		0.05	