WRITING THE KPC SURVEY REPORT

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

This updated version of "Writing the Survey Report" provides recommendations on the format and content of the KPC survey report.

I. FORMAT

It is useful to present material in the following manner:

A. Cover Page (include title, date, PVO/country, partner organizations, and author names)
B. Acknowledgments (list all supervisors, interviewers, etc., and their titles)
C. Table of Contents
D. Executive Summary (written last)
E. Background
F. Process and Partnership Building
G. Methods
H. Results
I. Discussion
J. Bibliography
K. Appendices

II. CONTENT

Background

This section of the report includes 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, economic conditions within the project area
- national standards/policies regarding maternal and child health (e.g., immunization, breastfeeding, or Safe Motherhood policies)

The author(s) should also give an overview of the project, namely the following:

- goals
- objectives
- intervention activities
- objectives of the KPC survey

Process and Partnership Building

The KPC was initially 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 utilization 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

Why Involve Local Partners/Stakeholders?

Partnership building is a dynamic and sometimes, an intensive process. When planning a KPC survey, it is important to consider the amount of time and resources needed to work with partners and build local capacity. The inclusion of local partners involves a great deal of boundary setting and negotiation between the various parties involved. Nevertheless, a commitment to partnership building in both research and project implementation is a means of consensus building and collaborative decisionmaking. This also ensures that the interests and concerns of all stakeholders are taken into account.

Methods

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

- Questionnaire
- KPC indicators
- Sampling design
- Training
- Data collection
- Data analysis

Questionnaire:

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

Study indicators:

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

The *Rapid CATCH* yields indicators that are useful to all CS projects, regardless of their intervention mix. Reporting on these core indicators is strongly encouraged by the CORE Monitoring and Evaluation Working Group. Projects are encouraged to consult the questions in the KPC2000+ modules to obtain supplemental information that is relevant to their specific project activities. In adapting the questionnaire, projects are reminded that the KPC should yield a concise and manageable set of indicators for reporting results. Additional information may be useful in developing project activities or strategies.

Sampling design:

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

While 30-cluster sampling has been traditionally used with KPC surveys, 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 clusters (if cluster sampling was used), households, and respondents; c) the number of clusters; and d) the number of interviews conducted within each cluster. It is also helpful to state any selection protocols employed during the study (e.g., what to do when there is more than one mother with a child less than two within the same household, or when the selected mother has more than one child less than two years).

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.

Data Analysis

- Method of data analysis (i.e., hand tabulation 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 very 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 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 may 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 age 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 very beginning of the Results section that lists all indicators, their numerators, denominators, percents, and confidence limits. The following table is an illustration.

INDICATOR	NUMERATOR	DENOMINATOR	PERCENT	CONFIDENCE LIMITS
Percentage of children age 0–23 months who are underweight (-2 SD from the median weight-for-age, according to the WHO/NCHS reference population)				

Percentage of children age 0–23 months who were born at least 24 months after the previous surviving child Percentage of children age 0–23 months whose births were attended by skilled health personnel Percentage of mothers with children age 0–23 months who received at least two tetanus toxoid injections before the birth of their youngest child Percentage of children age 0–5 months who were exclusively breastfed during the last 24 hours Percentage of children age 6–9 months who received breastmilk and complementary foods during the last 24 hours Percentage of children age 12–23 months who are fully vaccinated (against the five vaccine-preventable diseases) before the first birthday Percentage of children age 12–23 months who received a measles vaccine Percentage of children age 0–23 months who slept under an insecticide-treated net (in malaria risk areas) the previous night Percentage of mothers with children age 0–23 months who cite at least two known ways of reducing the risk of HIV infection Percentage of mothers with children age 0–23 months who report that they wash their hands with soap/ash before food preparation, before feeding children, after defection, and after attending to a child who has defecated Percentage of mothers of children age 0–23 months who know at least two signs of childhood illness that indicate the need for treatment Percentage of sick children age 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% of infants will be fully immunized in their first year of life.

Indicator: At the end of the project, 80% of children 12–23 months will be fully immunized.

KPC result: 65% of children 12–23 months of age in the survey are fully immunized. The confidence limits are calculated as plus or minus 10%.

Conclusion: We are 95% confident that the true proportion of fully immunized children in the population is between 55% and 75% (65% ±10%). The best estimate of the true proportion is 65%.

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% (the probability that 80% is the true proportion in the population, given the survey finding, is less than 5%).

The project should study other EPI findings from the survey to identify the barriers to achieving the objective. 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% of infants will be fully immunized in their first year of life.

Indicator: At the end of the project, 80% of children 12–23 months will be fully immunized.

KPC result: 76% of children 12–23 months of age in the survey are fully immunized. The confidence limits are calculated as plus or minus 9%.

Conclusion: We are 95% confident that the true proportion of fully immunized children in the population is between 67% and 85% (76% $\pm$ 9%). The best estimate of the true proportion is 76%.

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. That is, there is no evidence that the true proportion of fully immunized children in the population is less than 80% (the probability that 80% is the true proportion in the population, given the survey finding, is greater than 5%).

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% coverage. The project should study the 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).

Confidence Limit Formulas

Confidence limits with a simple random sample (SRS)

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

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

Z = 95% 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% confident that the true proportion in the population is between 33% and 47%. The best estimate for the true proportion in the population is 40%.

Table 1. Confidence Limits for a Simple Random Sample:

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

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 be representative of the entire population under study than a sample of randomly selected individuals. The implication of this bias is that the confidence limits of a finding from a cluster survey is 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 cluster (cluster id) the respondent belongs to, in order 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:

$P = p \pm z\sqrt{pq/n_e}$, **where** n_e = the ***effective sample size*** of the sample or sub-sample

Effective Sample Size (n_e) = **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 simple random sample. The design effect is used to correct the value of **n** used to calculate the confidence limit of a cluster survey.

Design Effect—In order 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 limit of a finding with the following methods:

1. First, 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). Second, 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). Third, report both confidence limits as the range of possible values.

2. 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.

3. 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

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

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

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

Discussion

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

1. relate key findings from the KPC survey to data from other sources
2. discuss the programmatic implications of the survey findings
3. 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:

- Demographic and Health Survey data for the country in which you are working
- Other local surveys
- Ministry of Health (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. 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 in which capacity
- Training schedule for supervisors and interviewers
- Project resource requirements of the survey, namely monetary costs and amount of staff time devoted to KPC activities