

RHRC CONSORTIUM MONITORING AND EVALUATION TOOLKIT

COMMUNITY-BASED SURVEY PROTOCOL

Purpose A community-based survey is used to determine population-based measurements and collect quantitative data about reproductive health.

Description In a community-based survey, a sample of all community members is interviewed using a structured, pre-coded questionnaire. As long as the sample is chosen following the rules of random sampling, the survey results represent the entire community even though only some people are interviewed. If the survey is of a sub-group in the population, such as women of reproductive age or adolescent men, then the results represent only that sub-group and not the community as a whole.

Ethical Considerations

Approvals

Before you begin a study, you must obtain certain permissions for ethical, political and logistical reasons. Some groups you may need to obtain permission from include UNHCR, Ministry of Health, civil authorities in your district, community representatives, your own organization and partner organizations. The consent of the individuals interviewed is always required.

Informed Consent

Every respondent has the right to refuse the interview, or to refuse to answer specific survey questions. The interviewers must respect this right.

Privacy

It is important that the interview with each respondent be conducted in a manner that is comfortable for her or him, and in which she or he is able to speak openly and honestly.

Each respondent will be contacted at home, and it is expected that the interview will take place there. No other adult man, woman or older child should be present or be able to hear the interview. Small children may be present. If the respondent indicates that she or he is uncomfortable holding the interview at home, the interview can be done at another place or at a later time.

Confidentiality

The interviewers may not discuss the respondents' answers with anyone, except the supervisor when clarification is needed. Respondents' names or other identifying information should not appear on the questionnaire; there should be no way to link a specific questionnaire to a specific respondent.

Data Collection Forms

A closed-ended, pre-coded questionnaire can be used. The same questionnaire can be used for both women and men respondents (questions specific to women are skipped for the men). The draft questionnaire should be pre-tested and

revised at least once; a second round of pre-testing and revision should be done if substantial changes are made after the first round.

It is best if the questionnaire is written in the locally-spoken language. If this is impossible, supervisors and interviewers should agree on the specific wording to be used in the local language as part of the interviewer training.

Sampling Plan

There are many considerations that come into play when designing a research study. Compromises are always being made on decisions such as sample size, acceptable error levels and sources of bias, based on the availability of resources (time, money, personnel). There is no one right way to decide what is acceptable. The following method for determining sample size is a conservative approach, and so should work in many situations.

The formula for calculating sample size is:

$$N = \frac{Z^2(p)(q)}{d^2}$$

- **Z** refers to the confidence limits of the survey results. If you would like to be 95% confident in your results, $Z=1.96$.
- **p** refers to the proportion of the population with the attribute you are looking for. $q=(1-p)$ If we had the data already, p and q would be very easy to calculate. Since we must estimate the proportions, we should be as conservative as possible. It is safest to maximize the variation, by assuming a 50/50 split in responses across questions.
- **d** refers to the desired precision of the estimate (for example, your results are accurate within a range of plus or minus 5%).

So, using these figures in the equation above, we get:

$$N = \frac{(1.96^2)(.5)(.5)}{.05^2} = 384$$

Given this calculation, we find that we would need a sample size of 384 to be 95 percent confident of our overall results, within a range of plus or minus 5 percent. That is to say, $N = 384$ is the sample size to use if we want observed percentages to be within plus or minus 5 percent of the unknown population parameter. This means our results will be a true reflection of the population 19 out of 20 times, although in a given instance, we don't know whether we are truly one of the 19/20 or if we are that 1/20 exception.

You can see that these numbers do not change based on population. If you wanted to always have a sampling error of plus or minus 5 percent and always have the most conservative estimate of your response distribution (that is, assume a 50/50 split in responses to your questions), your sample would always be 384. For populations under 10,000, this formula does not apply and you should consult the "Handbook for Family Planning Operations Research Design," Second Edition, by Andrew A. Fisher, John E. Laing, John E. Stoeckel, and John W. Townsend. It is published by the Population Council, who can be contacted at

www.popcouncil.org. Another statistics textbook could be consulted for sampling strategies for populations under 10,000, or you could consult with a statistician.

If a simple random sample of the population is not feasible, cluster sampling may be used. [Click here](#) for an explanation of cluster sampling.

If your sampling units vary considerable by size, you may want to use Probability Proportional to Size (PPS) sampling. [Click here](#) for an explanation of PPS sampling.

A table such as this may be useful for planning your survey. This table is from a cluster sampling plan.

Baseline Survey Sampling Frame <i>Community-based cluster sample</i>		
	Group A <i>(e.g., females age 15-44)</i>	Group B <i>(e.g., males age 15-59)</i>
Sample size		
Number of clusters		
Number of interviews per cluster		
Expected number of field supervisors		
Expected number of interviewers		
Expected number of days to complete survey field work		

Interviewer and Supervisor Characteristics

The ideal interviewers:

- are close in age to the respondents,
- have good communication skills in the local languages,
- are comfortable with discussing reproductive health topics, and are open to learning about reproductive health,
- will respect the dignity of respondents and confidentiality required by the survey,
- are interested and motivated to work,
- will reflect well on your organization in the field,
- have technical skills or training (depending on your subject),
- Previous experience with surveys is desired but not required.

Field supervisors should also exhibit these characteristics, have previous survey and reproductive health experience and be highly organized. Their responsibilities are to ensure that the sampling plan is followed, ensure that the interviewers are administering the questionnaire correctly, review every questionnaire in the field to ensure it is correctly and thoroughly filled out, code

the questionnaires, resolve interviewers' questions and respond to community questions or concerns.

Supervisors can be drawn from your organization and/or Ministry of Health staff. It will be extremely valuable for supervisors to be involved in the pre-tests and revisions of the questionnaire and interviewer training.

Interviewer Training

Interviewers should receive 3-5 days of training conducted by the field supervisors. The training emphasizes practice with the survey instrument so that the interviewers are entirely familiar with it and are comfortable asking the reproductive health questions. Training should include basic reproductive health technical information, so the interviewers can communicate with the respondents on the topic. During training, each interviewer completes at least 6 practice interviews.

A sample training plan is included below. [Click here](#) for an outline of topics for interviewer training used in a domestic violence survey among refugees in Lebanon.

Baseline Survey Interviewer Training Plan – Draft			
<p>Learning objectives: By the end of the 3-day training course, interviewers will demonstrate</p> <ul style="list-style-type: none"> ▪ Skill in administering the questionnaire for the baseline survey ▪ Ease with the reproductive health topics covered in the survey ▪ Understanding of their responsibilities in the field. 			
	Day 1	Day 2	Day 3
Morning	<ul style="list-style-type: none"> • Introductions • Orientation to NGO and project • Objectives of training and survey • RH Basics (Family Planning, Safe Motherhood, STIs/HIV, Gender-based Violence) 	<ul style="list-style-type: none"> • RH review • Translation of questionnaire • Field practice (3 interviews per interviewer) 	<ul style="list-style-type: none"> • RH review • Review field practice, finalize translation • Description of sampling plan and field procedures and interviewers' responsibilities
Lunch			
Afternoon	<ul style="list-style-type: none"> • Survey Basics • Practice interviews (with each other) 	<ul style="list-style-type: none"> • Review field practice, revise translation • Field practice 3 interviews per interviewer) in afternoon and evening 	<ul style="list-style-type: none"> • Logistics • Additional field practice (if needed) • Evaluation of training/ Assessment of interviewers

Time Frame To be determined by your organization. Community-based surveys are very time consuming, so organizations must make the time commitment from the start of the process. From the beginning of the survey planning to completion of the analysis and report can easily fill 4-6 months, especially if the staff have other

responsibilities. The field work itself (interviews) should be completed within 2-4 weeks.

[Click here](#) for a survey workplan format and example.

Analysis Plan

The data for quantitative analysis can be processed using any database or survey software, such as Epi Info for Windows, Epi Info 6, SPSS, SAS or Microsoft Access. (The advantage of Epi Info is that it can be downloaded for free from the cdc.gov web site <http://www.cdc.gov/epiinfo/>. Before each use of Epi Info, check the website for updates.)

A data entry program should be created using, for example, Epi Info for Windows, prior to the beginning of the fieldwork. Ideally, data will be entered as the questionnaires are completed, so that data entry will be complete shortly after the field work is complete.

A word to the wise – it is very important to save back-up copies of data on removable media such as floppy disks, zip disks or CD-roms regularly.

After data entry is complete, the data will be cleaned and checked in Epi Info. This can be done by running List and Frequencies commands to search for problems, such as duplicate or missing ID numbers and unallowed values (which should be few since the data entry program permits you to set minimum and maximum allowable values). Crosstabulations (“Tables” command in Epi Info) can also be used to check for unexpected variations among interviewers, supervisors or clusters.

It is common for the major part of the analysis to consist of frequencies and crosstabulations. Regression (to identify the main explanatory variables for knowledge, attitudes and behavior) also may be undertaken.

The initial task in the analysis will be to produce the draft dummy tables, which are the blank tables you prepare to fill in with survey findings when reporting your results. [Click here](#) to see sample “dummy tables”. Separately, for the samples of different populations (i.e., women and men), 3 steps will be required:

1. Recoding the variables with multiple responses, so that they have 2, 3 or, if absolutely necessary, 4 categories of responses. Recoding will be required for several of the socio-demographic characteristics (e.g., recoding age from individual age in years to 3 or 4 age groups; also required for education, occupation, marital status and others), and may also be needed for some of the other questions with multiple pre-coded answers (e.g., site of last delivery, source of contraceptives).
2. Running the frequencies for all of the variables listed in the dummy tables, including the recoded variables. This will provide the number and proportion of respondents answering each of the pre-coded (or recoded) responses (such as the proportion who spontaneously report that AIDS is transmitted by sexual intercourse).

3. Creating new summary variables for combining existing variables to know, for example, the proportion of respondents who know at least 2 correct ways AIDS can be transmitted, the proportion of respondents who know at least 3 contraceptive methods or the mean number of contraceptive methods known.

It is the decision of the researcher to determine how they analyze their data.

Use of Data Survey data can be used to establish baseline measures for project objectives or other key desired results, and/or to better understand the socio-demographic and program-related factors associated with different aspects of reproductive health. Later in a project, findings from a repeat survey can be compared to baseline levels.

Dissemination

Summary and detailed reports should be written. [Click here](#) for an outline of a survey report. The detailed report is most likely to be of interest to your organization and the national and provincial Ministry of Health and perhaps some of the NGOs working on similar topics. The summary report will have a broader distribution to other NGOs and Ministry offices in other provinces.

In addition to the written reports, you should present the survey results to communities as part of its education activities.

Additional Resources

Population Council: *The Handbook for Family Planning Operations Research Design*. Free of cost. Can be ordered online at <https://www.popcouncil.org/publications/puborder.html>, or from Population Council Office of Publications, One Dag Hammarskjold Plaza, New York, NY 10017-2201, USA.

Family Health International: [*Behavioral Surveillance Surveys: Guidelines for Repeated Behavioral Surveys in Populations at Risk of HIV.*](#)

Cluster Sampling

If a simple random sample of the population is not feasible, cluster sampling may be used. In each target area, 30 clusters will be selected randomly. In each cluster, 13 individuals (384/30) will be interviewed to attain the desired sample size in each area. If you are interviewing more than one population (ie men and women), you should interview 13 men and 13 women in each cluster. For most purposes, this sample of 30 clusters of 13 people is satisfactory.

In each cluster, the interview team will stand at the center of the cluster and spin a bottle (or use another random technique) to determine the direction of movement. Different directions should be selected for different samples. A random number will be chosen to identify the 1st household to select in each direction.

All eligible respondents in selected households will be interviewed. If an eligible respondent is not at home, he or she should be called if possible, or an appointment should be made by the interviewer to return later in the day. One attempt should be made to interview those who are not present. If the eligible person is still not present at the time of the appointment, he or she will be recorded as “Not Present.”

When all eligible respondents in a household have been interviewed (or appointments made to return later in the day), the interviewer and/or supervisor will select the next household by counting 5 houses to the right. If the border of the cluster is met before the desired number of interviews is reached, the interviewer and/or supervisor should return to the center of the cluster and choose a new direction and new random start.