Appendix: Guidelines for using mini-baseline survey and tally sheets

Sampling

Sampling means collecting data from a group in the population that is representative of the whole. It has been likened to eating a bowl of rice where you only have to try one spoonful to know if the food is good enough to eat[1].

Sample size

The public health survey is meant to give an idea as to what is happening in the community – it is not a statistically correct study. However the sample must be large enough for you to comfortably assume that it is fairly representative of the majority of the population and small enough not to waste resources collecting from too many people.

There are several ways of doing this.

1. Rough calculations based on population

For populations under 100 – 30-50 units[2]
For populations between 100-300 use 50-70 units
For populations between 300-1000 use 70-90 units
For populations over 1000 use 90-100 units

Units being households, groups or individuals

2. Using a computer programme

There are several computer programmes that calculate sample sizes.
http://www.surveysystem.com/sscalc.htm is a free website that allows you to calculate the sample size on screen. However, you need to understand the concepts of confidence intervals and confidence level.
EpiInfo can be downloaded from www.cdc.gov and has a section for finding sample size. Again, a certain degree of familiarity with statistical terms is required.

Weighting

If the different community groups or villages are of different sizes, you need to "weight" the sample size – you need to make sure the number of people interviewed in each community or village is proportional to the size of that community. You do this using the following formula:

Number of people in the community X number of the sample required
---

[1] Bolt E & Cairncross S. *Sustainability of hygiene behaviour and the effectiveness of change interventions* Delft: IRC; 2004

[2] Dale R. *Evaluation Frameworks for Development Programmes and Projects*. New Delhi: Sage; 1998

Total number of population being studied

An example of a weighted sample size

The catchment area for the Fatumasi clinic has six villages with a total population of 3620 people.
You need 100 people for questionnaire interviews
Add another 10-15 questionnaires in case you find that some are not answered properly and have to be discarded
In village A, there are 345 people so you use the above formula

$$\frac{345 \times 115}{3620} = 11$$

Village B has 1072 people so again you calculate

$$\frac{1072 \times 115}{3620} = 34$$

Then you do the same for all the other villages

Once you have the sample size, you need to decide how to find the households or other units to be interviewed. There are a couple of ways:

1. **Cluster sampling**

Especially if communities are far apart or very big, cluster sampling is a good approach, as logistically it is easier. Give all communities (or sections in a camp) a number and then select about 5% of these by picking numbers out of a hat or using random tables. If communities are very different form each other (for example some are in the hills and some are by the sea) then you need to make sure you select from both. In the case of cluster sampling, the sample size should be 50% larger than when using simple random selection.

2. **Random selection of households**

An example of this is to choose a starting point in the community or camp – the community centre, a religious building or a water point. Spin a bottle on the ground to choose the direction. Count two houses and interview the third. When you reach the edge of the community or camp, spin the bottle again and keep counting and spinning until you have your sample. If one house is closed, chose the one next to it and then count two before the next interview.

3. **Numbered units and randomised sampling**

Occasionally you may work in an area or camp where all the houses are in rows and have numbers (Abou Shouk camp in Darfur is an example). Here you can randomly pick houses before you go to the field using random tables or picking numbers out of a hat.

Training of enumerators

When considering who to use as an enumerator, it is best to consider skills such as communication and the ability to get on with the community. They need to be able to speak the local language and it is good to have a gender mix. Sometimes young women are not able to talk to older men or men cannot interview women so this needs to be considered when recruiting.

Never assume that everyone knows how to ask questions using a questionnaire. It is worth taking the time to do at least a days training on questionnaires, asking questions, not using positive or negative body language. It is also good to use the time to get the questionnaire translated into the local language (the enumerators can do this) and to pre-test it. This gives everyone a chance to familiarise themselves with the tool.
Chose one person to be the supervisor for each group and make sure they understand the importance of checking questionnaires in the evening for mistakes.

If you don't have time to train the enumerators, you should not think about doing a survey!

Translation of questionnaire

Once translated into the local language, it is good to "back translate" it into English to check for consistency. Sometimes terms and concepts lose their meaning in translation.

Adjusting the questionnaire to the current situation

You must adjust the questionnaire to suit your own situation. It is designed to cut and paste. If no one has toilets because they have recently arrived and are all living in a field then do not ask questions about toilets, ask about defecation sites instead.

Adjusting the tally sheets

Adjust the tally sheets when you adjust the questionnaire – this saves time later.

Coding and recording

When the enumerators return every day with completed questionnaires, take time to run through them and check for mistakes and "missing values" (blanks where there should be answers). Don't wait until the end. If one enumerator is consistently forgetting to ask questions or makes mistakes, you want to pick it up early.
If possible, enter the data onto the tally sheets every evening – you do not have to do this yourself but you do need to train the person doing it. Use the five-line method and then add totals.

Use of data

The important rule with surveys is:

If you are not going to use the data don't collect it!

Time allocation

Surveys always take longer than you think. Good planning means that there is a better chance of collecting useful information that will feed into the programme monitoring system.

Activity	Time allocated	Resources required
Prepare plan and budget	One hour	Transport Printing costs for questionnaire Stationary

		Lunch allowances or per diems Salaries
Prepare questionnaire and tally sheet Print questionnaires Work out sample size Decide on sample method	One day	Paper and a photocopier
Recruit enumerators	Two days – a week	Payment for advertisements Transport costs for interviewees
Train enumerators Translation Pre-testing Back translation	Two days	Training materials Lunch and snacks Transport for field visits Interpreter
Data collection Each questionnaire should take 30 minutes Each enumerator can do 10 questionnaires a day (minimum)	Two or three days Depending on distances	10 enumerators for one day or five enumerators for two days Two days for contingency Transport Lunch or snacks for enumerators and driver Clipboards and pens
Data recording	Two or three days	Designated data recorders Tally sheets Final tally sheet
Data analysis Report writing	Two days	The team leader The supervisors Pre-prepared tables
Total time	14 days minimum	