14.2: The questions

Quantitative data may be collected in field trials by a series of questions asked of the respondents that are compiled into a questionnaire. Additional quantitative data may be obtained by direct observation (for example, of what the house’s roof is made of or of whether a male has been circumcised), measurement (for example, weight), or after taking a tissue sample (for example, haemoglobin level). This section will cover issues related to data that are collected through questions.

2.1 Relation to study objectives, content, and duration

The questions to be included in a questionnaire should be developed to relate directly to the objectives of the study. Usually, at least an outline questionnaire will be drawn up in parallel with the formulation of the protocol for the trial. Most grant review committees expect to see such an outline in the trial funding application.

Questionnaires must be realistic, both in terms of content and length. For example, it may not be possible to obtain valid data on highly sensitive questions such as illegal or stigmatized behaviours through a structured questionnaire.

When a questionnaire survey is being planned, it will often seem attractive to add questions that do not relate directly to the objectives of the study but which may be of interest for other reasons (see also Chapter 12, Section 2.1). As a general rule, this temptation should be resisted, as lengthening interview schedules is likely to lead to a higher non-participation rate, and time devoted to questions of peripheral interest may be at the expense of time on more important questions, with a consequent lowering of the quality of the information collected on the latter. It is good practice to go through a draft questionnaire, specifying which objective or important trial outcome each question will contribute to, with the aim of deleting any which cannot be clearly justified on these grounds. Nonetheless, in some circumstances, it may be desirable to ask other questions if this increases the likelihood of participation in the survey or serves to divert attention away from the main questions, in order to reduce the chance of biased responses. For example, it may be
more acceptable to ask questions about sexual behaviour in the context of a more general behavioural survey than to include only questions that concern sexual behaviour. Similarly, if particular adverse effects are expected from an intervention, it will usually be best to also include questions about effects thought to be unrelated to the intervention, as this may help identify any biases in response between intervention and control groups that are not directly attributable to the intervention.

Few respondents will be willing to complete a questionnaire that takes more than 20–30 minutes, and, even if they do, the quality of responses may well decline if the respondent gets bored or tired. In general, it is best if a questionnaire can be kept to less than 30 minutes, though this can sometimes be extended if it includes a variety of different activities, such as answering questions about photographs or scenarios or taking physical measurements, rather than only questions and answers.

2.2 Development of questions

A plan for the development of the questions to be included in a questionnaire survey is given in Box 14.1, and Box 14.2 gives a checklist of points that should be considered in drafting questions.

Increasingly, standardized questionnaires are being developed and shared. These draw on questionnaires and interviews that have been conducted in many countries and studies, and often the questions and responses have been translated into many different languages. An example of this is the Economic and Social Research Council (ESRC) question bank (<www.surveynet.ac.uk/sqb>), which has hundreds of survey questionnaires in it. With the advent of standards for data documentation (see <http://www.ddialliance.org>), searching and browsing for questions on particular themes will be easier and more extensive. While it is unlikely that complete questionnaires can be copied for new trials, it is important to utilize the resources and knowledge from previous studies to avoid making the same mistakes and to build on existing knowledge.

Box 14.1 Checklist for the development of a questionnaire

1. Define the information that is required from the questions. Some items of information may only require a single question, such as name or sex, while others require a series of questions such as socio-economic status or episodes of illness in the past week.

2. Formulate draft questions. Attention to the wording of questions is important, as slight variations may result in different responses. For example, ‘Where do you normally seek help when your child has diarrhoea?’ vs ‘Where did you seek help when your child last had diarrhoea?’. Box 14.2 gives a checklist of points that should be considered in drafting questions. In general, it is a good idea to search for, and to critically review, how others have asked specific questions, especially if these questions have been formally validated.

3. Informally test the questions. This may involve trying them out on different members of the study team and discussing them with those knowledgeable of the study area, including residents. It may be necessary to base someone in the community under study (ideally, someone with anthropological or social science skills) to investigate how different questions will be perceived to find out if there are taboos regarding certain topics, if there are local words for some illnesses or conditions, and the extent to which these correspond to the investigator’s definitions (for example, many communities have special words for measles, night blindness, sexual intercourse, depression, or lethargy).

The investment warranted for such qualitative studies will depend upon local sensitivities regarding the items on which information is required and the degree to which each question is critical for the trial. For example, it will require less work to find out how to ask...
questions about breastfeeding practices than to formulate appropriate questions on aspects of sexual behaviour. As a result of such investigations, the original draft questions may have to be modified. Some may even have to be abandoned if research indicates that valid information is unlikely to be elicited through a questionnaire survey.

4. Prepare a first draft of the questionnaire for pilot testing.

5. Translate each question into the language(s) of the study population, followed by independent back-translation by someone who does not know the original questions, with reconciliation of any discrepancies—ideally followed by further independent translation and back-translation (see Section 2.5).

6. Prepare a draft instruction manual for interviewers and their supervisors (see Section 4.4).

7. Pilot-test the questionnaire in field conditions, preferably in an area adjacent to the study area and using the interviewers who will work on the main survey (see Chapter 13).

8. Analyse the experience in the pilot test and the data collected.

9. Reformulate the questionnaire, with further translation and back-translation of any amended questions, followed by further pilot testing, especially if important changes have been made to questions related to primary or secondary trial outcomes.

10. Finalize the questionnaire for the main survey, along with the instruction manual for interviewers (see Section 4.4).

Box 14.2 Checklist of points to consider when drafting questions

1. Keep wording informal, conversational, and simple. Avoid words longer than three or four syllables.

2. Avoid jargon and sophisticated language; assessing understanding at the pre-test and pilot test stages is essential. The wording of all questions must be appropriate to the educational, social, and cultural background of the respondents.

3. Check the cultural relevance to the respondents of concepts used. Ensure mutual understanding between the interviewers and the respondents, paying attention to cultural and educational differences.

4. Avoid long questions, but vary the length of questions to avoid administration of the questionnaire becoming repetitive and boring for the interviewer or interviewee.

5. It may be necessary to define a term or a concept before asking about it. If the definition is short, it can be included in the question, but otherwise it is better given separately before the question is asked.

6. Avoid leading questions that may bias the respondent to a particular answer (for example, ‘Do you think the improved clinic arrangements are better?’).

7. Avoid open questions beginning ‘Why?’.

8. Avoid negative questions (for example, ‘Do you not think . . . ’—in some cultures, the answer ‘no’ indicates ‘I do not think . . . ’; in other cultures, the answer ‘yes’ indicates ‘Yes, I do not think . . . ’).

9. Where possible, avoid hypothetical questions, as some respondents will find these difficult to answer (for example, ‘If the bus fare was less, would you come to the clinic more often?’).

10. Keep to a single subject for each question. For example, do not say ‘Do the cost and times of the clinic prevent you going?’.

11. Pay particular attention to sensitive issues. Review the inclusion of very sensitive ones. If they are to be retained, pay very careful attention to the wording, and consider the use of indirect approaches. Think carefully about their position within the questionnaire (see Section 3.2).

12. Check the adequacy of the lists of responses to ‘closed’ (see Section 2.3.2) questions. For example, ensure a food list covers most things normally eaten in the community concerned. It is usually a good idea to include an ‘other (specify)’ category, unless you are sure that every possible answer is in the list (such as male and female for gender). But it is also important that only
a relatively small proportion of responses (definitely less than 10%) end up being in the ‘other (specify)’ category. This should be checked in the pilot test, with additional categories being added for the commoner responses that were initially in the ‘other (specify)’ category.

13. Never include an ‘other’ category without asking the respondent to specify what the response was—as in ‘other (specify)’—and leave space for the respondent or interviewer to write the specific answer next to this code.

2.3 Types of question

Information may be sought on opinions or facts through a questionnaire. The distinction between the two is not always clear, but, in general, the collection of data on the latter is easier to plan. Local sensitivities will influence the reliability with which either kind of information may be obtained. For example, in some cultures, it is considered unlucky to count your children, so asking a parent ‘How many children do you have?’ may be too direct an approach.

2.3.1 Historical recall

Information may be sought about the present (for example, ‘Does your child have fever now?’) or about the past (‘Did your child have an episode of fever in the last month?’). The advantage of asking about the present situation is that responses are not susceptible to memory lapses, and furthermore they will usually be more amenable to validation (see Section 2.4). The reliability of historical information decreases the further back in time the question relates to, and is influenced greatly by the importance of the event to the person (also referred to as its salience). Thus, deaths will be remembered better than hospital admissions, which, in turn, will be remembered better than illness episodes not requiring hospital admission. To obtain reliable information on mild, or even moderately severe, fevers, diarrhoea, or respiratory infections, the recall period probably should not exceed a week. The implication of this for longitudinal studies in which these outcomes are of interest, is that at least weekly surveys of the study group will be necessary to collect reliable information.

2.3.2 Open and closed questions

A ‘closed’ question is one that allows only a defined set of answers which have been anticipated and categorized in advance (for example, ‘Do you own a radio?’ 0 = No; 1 = Yes). Replies to an ‘open’ question can take any form and should, whenever possible, be recorded in the respondent’s own words (for example, ‘What were the symptoms your child had before being taken to the health facility?’). It is possible to ask a question with a closed list of responses in an open way, with the answer being assigned to one of a previously compiled list of codes held by the interviewer (for example, ‘What did you eat yesterday?’, with a list of types of food on the questionnaire for the interviewer to tick off those mentioned). This may produce a different response from asking closed questions about each of the items on the list. Reading out the list will remind the respondent of the possibilities but may also tend to produce affirmative answers as a gesture to ‘please’ the interviewer or because the respondent is embarrassed to admit that they have not eaten a high-status food such as meat. If the information is sufficiently important, both approaches can be used, the list of unmentioned possible answers being read out after initial responses are recorded without such prompting. The two responses should be recorded separately. For example, against each category, there could be three options: ‘0 = No’; ‘1 = Yes, unprompted’; ‘2 = Yes, prompted’. An analogy is medical history taking where questions about specific signs and symptoms might be asked after an initial neutral enquiry such as ‘What is the problem today?’.
In preliminary qualitative investigations, open questions are likely to be preferred to determine the full range of possible responses. As a general rule, however, for questions that are to be administered in a large survey, closed questions are better, as it is very tedious and time-consuming to go back to code the open answers subsequently. It is important that they are the ‘right’ closed questions, of course. This requires careful research and the avoidance of the premature administration of a questionnaire that may be simple to administer, code, and analyse but which does not provide the information required to meet the study objectives.

2.4 Validation

The principles underlying the validation of a questionnaire are similar to those for validating a diagnostic test. The objective is to determine to what extent the answers given to a question correspond to the ‘true’ situation. Problems arise if there is no independent way of ascertaining what is ‘true’. If a mother is asked ‘Does your child have fever now?’ the temperature of a child might be measured independently, and the response to the question validated against the direct measurement (by defining temperatures above some limit as ‘fever’). It will usually be impossible to validate the responses to a question such as ‘Did your child have fever yesterday?’ If a ‘gold standard’ exists, i.e. a means of obtaining an independent measure of the true response, the sensitivity and specificity of a given question can be assessed. The sensitivity of the question is the proportion of true positive responses that are reported as positive (for example, the proportion of all children with a current fever who are reported as having fever by their mother). The specificity is the proportion of true negative responses where the question produces a ‘negative’ response (for example, the proportion of all children without a current fever who are classified as not having a fever by questioning their mother).

The relative importance of sensitivity and specificity in intervention trials is discussed in Chapter 12, Section 4.2.

If there is no ‘gold standard’, other characteristics of the responses to questions must be evaluated to assess their usefulness in a particular survey. A minimal requirement for a question should be that the respondent gives the same answer to the same question at different times if the circumstances have not changed (i.e. responses should be ‘repeatable’). Also, if different interviewers administer the same question to the same person, the same answer should be obtained (i.e. responses should be ‘reproducible’). Repeatability and reproducibility are not a guarantee of validity, of course. The question ‘Do you beat your spouse?’ might be answered consistently over time to the same interviewer and reproducibly to different interviewers, but it may still be a very poor way of detecting spouse beaters! Also, a man might consistently report that his wife is his only sexual partner, even if this is not the case.

If a question fails to induce consistent answers, either within or between interviewers, it may be because of a fault in the question or in the interviewers or be due to the respondent deliberately varying their responses for some reason. For example, the respondent might reason that, ‘if I am being asked the same question a second time, this must be because they didn’t like my first answer, so I’d better change it’!

2.5 Translation

It will often be necessary to translate the questionnaire into local languages. Such translation should be undertaken with care and attention to detail, as it is easy for the sense of a question to be changed, sometimes substantially, by the translation process. For example, apparent differences in responses to a question asked to those in different language groups may be due entirely to variations in the translation processes. Words for some illnesses or concepts may not
exist in a language, and this may necessitate major changes in the wording of questions. An apparently equivalent word may exist, but it may be used in a different way and cover a narrower, or wider, range of conditions. For example, there may be several local words used to describe acute respiratory infections, one of which corresponds closely to what we mean by pneumonia. Conversely, difficulties may be encountered when one local word is used to encompass several different conditions. In studies of meningitis in The Gambia, for example, there was difficulty in finding terms to distinguish between a ‘floppy’ and a ‘stiff ’ neck (B. Greenwood, personal communication).

Once a questionnaire has been translated into a local language, it should be independently back-translated into the original language. Comparing the original text with the back-translated text will indicate possible areas of confusion where attention to the original translation will be required.