National Institute for Health Research

The NIHR Research Design Service for the East Midlands

The NIHR Research Design Service for Yorkshire & the Humber

Surveys and Questionnaires

Authors

Nigel Mathers

Nick Fox

Amanda Hunn

Nigel Mathers   Academic Unit of Primary Medical Care
Community Sciences Centre
Northern General Hospital
Herries Road
Sheffield
S5 7AU

Nick Fox   School of Health and Related Research (ScHARR)
University of Sheffield
Regent Court, 30 Regent Street
Sheffield
S1 4DA

Amanda Hunn   Left COREC at the end of December 2006

Last updated: 2009

© Copyright of The NIHR RDS EM / YH (2009)

Table of Contents

1. Introduction

The survey is probably the most commonly used research design in health services research and the social sciences. We have all been asked to take part in a survey at some time. As consumers we are asked about our shopping habits, as users of services we are asked for our opinions of services.

The survey is a flexible research approach used to investigate a wide range of topics. Surveys often employ the questionnaire as a tool for data collection. This resource pack considers the use of surveys and questionnaires in health and social care research.

<div style="text-align:center">**LEARNING OBJECTIVES**</div>

Having successfully completed the work in this chapter, you will be able to:
1. Understand why you might want to use a **survey**.
2. Describe how to select a **sample** for a survey.
3. Understand why you might want to use a **questionnaire**.
4. Understand how the **method** used for data collection influences the design of the questionnaire.
5. Distinguish between a structured **questionnaire**, **semi-structured questionnaire** and **a topic guide**.
6. Design your own **questionnaire** and **coding frame**.
7. Distinguish between **open-ended** and **closed questions**.
8. List possible ways of increasing your **response rate**.

2. What is a survey?

Surveys are a very traditional way of conducting research. They are particularly useful for non-experimental descriptive designs that seek to describe reality. So, for instance, a survey approach may be used to establish the prevalence or incidence of a particular condition. Likewise, the survey approach is frequently used to collect information on attitudes and behaviour. Some issues are best addressed by classical experimental design where participants are randomised to either an **intervention** group or a **control** group. In the real world it is not always a very practical design. There may be good reasons, either ethical or practical, why participants cannot be randomly assigned to a particular intervention. It may also be impossible to identify a control group. Control over the randomisation process can also be difficult to achieve.

Surveys can take many forms. A survey of the entire population would be known as a **census**. However usually surveys are restricted to a representative sample of the potential group that the researcher is interested in, for reasons of practicality and cost-effectiveness. Most surveys take one of the following forms:

Cross-sectional Surveys

Surveys that are carried out at a just one point in time are known as a **cross-sectional** in design. They provide us with a snapshot of what is happening in that group at that particular time. They usually take a descriptive or exploratory form that simply sets out to describe behaviour or attitudes. So for example, if you want to measure some aspect of client satisfaction, then a cross-sectional descriptive survey would be the recommended approach. Likewise, if you wish to establish the prevalence of depression amongst new mothers, a postal survey might be an appropriate approach.

Longitudinal Surveys

Alternatively surveys can be **longitudinal**. A **longitudinal survey** rather than taking a snap-shot, paints a picture of events or attitudes over time. This may be a matter of months or years. There may be only two discrete surveys or there may be many repeated waves over a long period of time. Longitudinal surveys usually take one of two forms:

- **cohort surveys** - which follow the same group of individuals over time, or
- **trend surveys** - which take repeated samples of different people each time but always use the same core questions.

Cohort studies are particularly useful in tracking the progress of particular conditions over time, whereas trend studies set out to measure trends in public opinion and behaviour over time. For instance, take the client satisfaction survey which was mentioned earlier. If we wanted to compare levels of client satisfaction year on year, then a **longitudinal trend** survey would be recommended. With a trend study it is not necessary to interview the same individuals each time. In fact it is probably better to deliberately avoid the same people since the very fact of participating in a survey can raise levels of awareness and change behaviour. This is particularly true if you are trying to measure awareness of a promotion campaign. A particularly well known version of a trend survey is the General Household Survey which is carried out on an annual basis.

A cohort study is more difficult to carry out than a trend survey because the same individuals have to be traced over time and inevitably some participants move house, some fall ill and die, and some just refuse to participate. This loss of participants is known as 'attrition'. Sample size calculations are even more important for cohort surveys because high levels of attrition can result in too small a sample in

the final stages of the survey. Ideally expected levels of attrition should be calculated and allowed for in the initial sample selection. This means that the early stages of the cohort may be unnecessarily large but in turn this means that you will have adequate numbers in the final wave. A fine example of a cohort study is the National Child Development Study, which is based on an initial sample of children born in one week in 1947 and continued to follow them through over many years.

Explanatory or Correlational Surveys

We have described how some surveys seek only to describe events and attitudes. It is also possible for surveys to take an explanatory or correlational approach. This means that by using survey data the researcher would try to explore causal relationships between two or more variables. Demonstrating a causal relationship using survey data will always be more difficult than using an experimental design. Nevertheless there will always be situations in which an experimental design is just not possible. Using a longitudinal approach may also help in trying to identify a causal relationship. Statistical tests can be used to show statistically significant differences between groups in a survey. Confounding variables can also be controlled for in the data analysis.

2.1 What are the advantages of using a survey?

- *Surveys have internal and external validity* - A survey which is based on some form of random sampling technique will produce a sample which is representative of the particular population under study and will produce findings which may be generalised to the wider population. Randomised clinical trials (RCTs) on the other hand often have very stringent inclusion and exclusion criteria which make generalisations of the findings very difficult to apply in the real world.
- *Surveys are efficient* - Because surveys can use a random sampling technique to recruit participants, relatively small sample sizes can be used to generate findings which can be used to draw conclusions about the whole population. They are thus a very cost-effective way of finding out what people do, think and want.
- *Surveys can cover geographically spread samples* - Surveys can be undertaken using a wide range of techniques including postal questionnaires and telephone interviews. This means that participants who are widely dispersed can be accessed and included in the sample.
- *Surveys may have ethical advantages* - Since most surveys do not expose individuals to possibly invasive techniques or withhold treatment, they may be considered more ethical, since the individuals included in a study will merely be exposed to events that occur in the real world and would have taken place anyway.
- *Surveys are flexible* - Surveys can easily be combined with other methods to produce richer data. So for instance, you might want to consider also using diaries, focus groups, or in-depth interviews.

2.2 Limitations of the survey approach

- *Surveys are dependent upon the chosen sampling frame* –
  The representativeness of a survey is entirely dependent upon the accuracy of the sampling frame used. Sometimes it is not possible to identify an accurate or up-to-date sampling frame.
- *Surveys are not so good at explaining why people think or act as they do* - Surveys can tell us how many people behave in a certain way or how many patients were dissatisfied with their treatment, but they may be limited in the information they can provide as to **why** this is so (although asking open-ended questions can allow you to find out more). Qualitative research, such as focus groups, is usually much better at answering 'why' questions.
- *Interview surveys are only as good as the interviewers asking the questions* - The outcome of a survey may be influenced by interviewer error and bias. It is important that all interviewers receive proper training and are thoroughly briefed on each project. For the details on how to

approach Interviews you should refer to The NIHR RDS EM / YH Resource Pack: '*Using Interviews in a Research Project*'.

3. Methods of collecting survey data

It is important to remember that a survey is a type of research design. In contrast, an interview or a postal questionnaire is a method of data collection. There is a wide range of methods available for collecting data covering human participants, but the three main methods of collecting survey data are:

1. face-to-face interviews

2. telephone interviews

3. questionnaires

The selection of the appropriate method depends upon a number of factors, including:

- access to potential participants/respondents
- the literacy level of respondents
- the subject matter
- the motivation of the respondents
- resources

We will now cover each method in more detail.

3.1 Face-to-face interviews

Face-to-face or personal interviews are very labour intensive, but can be the best way of achieving high quality data. Face-to-face interviews are preferable:

- when the subject matter is very sensitive, but not personal
- if the questions to be coded are very complex or
- if the interview is likely to be lengthy

Face-to-face interviews can take both qualitative and quantitative approaches but surveys tend to take a quantitative approach. Interviewers carrying out face-to-face interviews for a quantitative study will use a highly structured interview schedule. Overall face-to-face interviews are more expensive than other methods but they can collect more complex information and are also useful when the subject matter is not of great personal interest to the respondent who would be unlikely to complete a postal questionnaire.

Face-to-face interviews are appropriate if you need to show anything such as diagrams or where certain disabilities may make completing a questionnaire in another fashion prohibitive. Do not conduct face-to-face interviews where you are asking personal questions such as about sexual history or drug taking, as the information you receive is likely to be inaccurate.

3.2 Telephone interviews

Telephone interviews can be a very effective and economical way of collecting quantitative data, if the individuals in the sampling frame can all equally be accessed via a telephone and if the questionnaire is fairly short. This may not be an appropriate method for a deprived population where telephone ownership is likely to be low, but can be ideally suited to a busy professional respondent, such as a social worker or a hospital consultant, if prior appointments are made. Telephone interviews are

particularly useful when the respondents to be interviewed are widely geographically distributed, but the complexity of the interview is limited without the use of visual aids and prompts. The length of a telephone interview is also limited, although this will vary with subject area and motivation. A prior appointment and covering letter may enhance the response rate and length of interview. However, as with the face-to-face method, once the person is committed they are more likely to complete the survey.

When it comes to data collection, telephone interviews are sometimes recorded using a tape-recorder or the answers can be typed directly into a computer as the interview is being conducted.

When doing a face-to-face or telephone survey of respondents in their own homes it is important to do some evening calling, otherwise the survey may be restricted to those who are at home during the day. For guidelines on conducting a face-to-face interview see The NIHR RDS EM / YH Resource Pack: '*Using Interviews in a Research Project*'.

3.3 Questionnaires

Questionnaires are a useful option to consider when conducting a postal survey. They can be cheaper than personal interviewing and quicker if the sample is large and widely dispersed. For any postal survey regardless of the sample size you must allow at least six weeks for the first wave of questionnaires to be returned, and another four weeks for each successive mailing. As with telephone interviewing, a postal survey is useful if your respondents are widely distributed. However, due to the lack of personal contact between the respondent and the researcher, the design and layout of the questionnaire is all important.

All mailed questionnaires should be accompanied by a covering letter and include a stamped, addressed envelope. In general, postal surveys tend to have lower response rates than face-to-face or telephone interviews. However questionnaires sent to populations with a covering letter from their general practitioner tend to have very high response rates.

As an alternative to mailing the questionnaire, it is possible to hand them out directly to your potential respondents in your chosen sampling frame. For instance you may decide that questionnaires can be handed out directly to parents with young children attending a nursery. Another example might be a health visitor visiting mothers six weeks after birth and asking them to complete a questionnaire. In both cases it is relatively easy to approach respondents in these circumstances and you are likely to achieve a much higher response rate than would be possible with a postal survey. The main drawback of this approach is that your captive audience may in some way be biased. For example, if you carry out a survey of client satisfaction which is restricted only to those patients attending one drop in centre, then the results will be biased towards the views of the most frequent attenders and consequently those people with most problems.

Questionnaires can be either devised by the researcher or they can be based upon some ready-made index. If you choose to design your own questionnaire for self completion, then the rules governing the style and layout are the same as those for designing a questionnaire for a postal survey (see above).

There are now many **pre-existing questionnaires** covering a wide range of conditions and therapy areas as well as quality of life instruments and satisfaction measures. Some of these are designed for self-completion, others are designed to be administered by an interviewer. There are obvious advantages to using such questionnaires, including the fact that many of these have already been well validated and tested for reliability, and there may well be normative data available as a baseline for you to compare your results with.

Many of these questionnaires are copyright protected and you may need the author's permission to use them. Likewise, using some questionnaires incurs a charge for each participant. You will therefore need to check whether or not there is a charge before you decide which questionnaire to use.

Remember there is no need to reinvent the wheel, so before designing your own questionnaire, you should spend time investigating what material exists already.

EXERCISE 1

1. What are the three most important advantages of a face-to-face interview survey?

2. What are the three main disadvantages of a face-to-face interview survey?

EXERCISE 2

1. What are the main advantages of a telephone survey?

2. What are the main disadvantages of a telephone survey?

EXERCISE 3

1. What are the main advantages of a postal survey?

2. What the main disadvantages of a postal survey?

Answers to exercises are given at the back of this resource pack.

4. Sampling for surveys

4.1 Why do we need to select a sample?

In some instances the sample for your study may be the same as the population under investigation. If the participants in your study are very rare, for instance a disease occurring only once in 100,000 children, then you might decide to study every case you can find. More usually, however, you are likely to find yourself in a situation where the potential participants in your study are much more common and you cannot practically include everybody.

So it is necessary to find some way of reducing the number of participants included in the survey without biasing the findings in any way. Random sampling is one way of achieving this, and with appropriate statistics such a study can yield valid findings at far lower cost. Samples can also be taken using non-random techniques, but in this chapter we will emphasise random sampling, which if conducted adequately, will ensure internal validity.

4.2 Random sampling

To obtain a random (or probability) sample, the first step is to define the population from which it is it be drawn. This population is known as the *sampling frame.* For instance, you are interested in doing a survey of children aged between two and ten years who have been adopted in the last year. Or you want to study adults (aged 16-65 years) diagnosed as having asthma and receiving drug treatment for asthma in the last six months, and living in a defined geographical region. In each case, these limits define the sampling frame.

The term *random* may imply to you that it is possible to take some sort of haphazard or *ad hoc* approach, for example stopping the first 20 people you meet in the street for inclusion in your study. This is not random in the true sense of the word. To be a 'random' sample, every individual in the population must have an *equal probability* of being selected. In order to carry out random sampling properly, strict procedures need to be adhered to.

Random sampling techniques can be split into *simple random sampling* and *systematic random sampling.*

Simple random sampling

If selections are made purely by chance this is known as simple random sampling. So, for instance, if we had a population containing 5000 people, we could allocate every individual a different number. If we wanted to achieve a sample size of 200, we could achieve this by pulling 200 of the 5000 numbers out of a hat. This is an example of simple random sampling. Another way of selecting the numbers would be to use a table of random numbers. These tables are usually to be found in the appendices of most statistical textbooks. However, these are only valid if the researcher is blind to the table.

Systematic random sampling

Systematic random sampling is a more commonly employed method. After numbers are allocated to everybody in the population frame, the first individual is picked using a random number table or out of a hat and then subsequent participants are selected using a fixed sampling interval, i.e. every nth person.

Assume, for example, that we wanted to carry out a survey of homeless people attending drop in centres in one Region. There may be too many to interview everyone, so we want to select a representative sample. If there are 3,000 homeless people attending the centres in total and we only require a sample of 200, we would need to:

- calculate the sampling interval by dividing 3,000 by 200 to give a sampling fraction of 15
- select a random number between one and 15 using a set of random number tables or numbers out of a hat
- if this number were 13, we select the individual allocated number 13 and then go on to select every 15th person after that

This will give us a total sample size of 200 as required.

Care needs to be taken when using a systematic random sampling method in case there is some bias in the way that lists of individuals are compiled, for example if all the husbands' names precede wives' names and the sampling interval is an even number, then we may end up selecting all women and no men.

Stratified random sampling

Stratified sampling is a way of ensuring that particular strata or categories of individuals are represented in the sampling process.

If, for example, we want to study consultation rates in a general practice, and we know that approximately four per cent of our population frame is made up of a particular ethnic minority group, there is a chance that with simple or systematic random sampling we could end up with no ethnic minorities (or a much reduced proportion) in our sample. If we wanted to ensure that our sample was representative of the population frame, then we would employ a stratified sampling method.

- First we would split the population into the different strata, in this case, separating out those individuals with the relevant ethnic background.
- We would then apply random sampling techniques to each of the two ethnic groups separately, using the same sampling interval in each group.
- This would ensure that the final sampling frame was representative of the minority group we wanted to include, on a pro-rata basis with the actual population.

Disproportionate sampling

Taking this example once more, if our objective was to compare the results of our minority group with the larger group, then we would have difficulty in doing so, using the proportionate stratified sampling just described, because the numbers achieved in the minority group, although pro-rata those of the population, would not be large enough to demonstrate statistical differences.

To compare the survey results of the minority individuals with those of the larger group, then it is necessary to use a disproportionate sampling method. With disproportionate sampling, the strata selected are not selected pro-rata to their size in the wider population. For instance, if we are interested in comparing the referral rates for particular minority groups with other larger groups, then it is necessary to over sample the smaller categories in order to achieve statistical power, that is, in order to be able to demonstrate statistically significant differences between groups.

Note that, if subsequently we wish to refer to the total sample as a whole, representative of the wider population, then it will become necessary to re-weight the categories back into the proportions in which they are represented in reality. For example, if we wanted to compare the views and satisfaction levels of women who gave birth at home compared with the majority of women who have

given birth in hospital, a systematic random sample, although representative of all women giving birth would not produce a sufficient number of women giving birth at home to be able to compare the results, unless the total sample was so big that it would take many years to collate. We would also end up interviewing more women than we needed who have given birth in hospital. In this case it would be necessary to over-sample or over-represent those women giving birth at home to have enough individuals in each group in order to compare them. We would therefore use disproportionate, stratified random sampling to select the sample in this instance.

The important thing to note about disproportionate sampling is that randomisation is still taking place within each stratum or category. So we would use systematic random selection to select a sample from the 'majority' group and the same process to select samples from the minority groups.

Cluster sampling

Cluster sampling is a method frequently employed in national surveys where it is uneconomic to carry out interviews with individuals scattered across the country. Cluster sampling allows individuals to be selected in geographical batches. So for instance, before selecting at random, the researcher may decide to focus on certain towns or electoral wards. Multi-stage sampling allows the individuals within the selected cluster units to then be selected at random.

Obviously care must be taken to ensure that the cluster units selected are generally representative of the population and are not strongly biased in any way. If, for example, all the general practices selected for a study were fund holding, this would not be representative of all general practice.

Note that even if the researcher randomly selects the initial clusters this does not constitute a truly random sampling method.

4.3 Non-random sampling

Non-random (or non-probability) sampling is not used very often in quantitative social research, but it is used increasingly in market research surveys and commissioned studies. The technique most commonly used is known as quota sampling.

Quota sampling

Quota sampling is a technique for sampling whereby the researcher decides in advance on certain key characteristics which s/he will use to stratify the sample. Interviewers are often set sample quotas in terms of age and sex. So, for example, with a sample of 200 people, they may decide that 50% should be male and 50% should be female; and 40% should be aged over 40 years and 60% aged 39 years or less. The difference with a stratified sample is that the *respondents in a quota sample are not randomly selected* within the strata. The respondents may be selected just because they are accessible to the interviewer. Because random sampling is not employed, it is not possible to apply inferential statistics and generalise the findings to a wider population.

Convenience or opportunistic sampling

Selecting respondents purely because they are easily accessible is known as *convenience sampling*. Whilst this technique is generally frowned upon by quantitative researchers, it is regarded as an acceptable approach when using a qualitative design, since generalisability is not a main aim of qualitative approaches.

EXERCISE 4

Read the descriptions below and decide what type of sample selection has taken place.

1. School children, some with foster parents and some with natural parents are identified from school records. Method: children are selected randomly within each of the two groups and the number of children in each group is representative of the total child population for this age group.

2. Children with and without chronic asthma are identified from GP records. Method: the children are selected so that exactly 50% have chronic asthma and 50% have no asthma. Within each group, the children are randomly selected.

3. A survey of the attitudes of mothers with children under one year. Method: interviewers stop likely looking women pushing prams in the street. The number of respondents who fall into different age bands and social classes is strictly controlled.

4. A survey of attitudes of drug users to rehabilitation services. Method: drug users are recruited by advertising in the local newspaper for potential respondents.

5. A postal survey of the attitudes of males to use of male contraceptives. Method: all male adults whose National Insurance numbers end in '5' are selected for a survey.

6. A study of the length of stay of patients at Anytown General Hospital. Method: all patients admitted to wards 3, 5, and 10 in a hospital are selected for a study.

This is an opportunity to review your learning on this first part of the resource pack. Read the extract from a journal article given below:

'National asthma survey reveals continuing morbidity'

(Prescriber, 19 March 1996 p. 15)

'A preliminary analysis of a survey of 44,177 people with asthma has revealed that for many the condition causes frequent symptom and substantially interferes with daily life. There is also a trend for older people with asthma to experience more problems. More information about treatment was seen by many as the best way to improve care.

The Impact of Asthma Survey was conducted by Gallup on behalf of the National Asthma Campaign with funding from Allen & Hanburys. Questionnaires were given to people with asthma via surgeries, pharmacies, retail outlets, the media and direct mailing in the autumn of 1995; the respondents were therefore self-selected and may not be representative of the population with asthma.

Asthma symptoms were experienced on most days or daily by 41 per cent of survey respondents, ranging from 18 per cent of the under-11s to 55 per cent of pensioners. Waking every night with wheeze, cough or breathlessness was reported by 13 per cent and 43 per cent say they are woken by symptoms at least once a week.

About 20 per cent consider that asthma dominates their life, ranging from 17 per cent in children to 37 per cent in the over-60s; over 40 per cent of each age group say the condition has a moderate impact on their quality of life.'

Now answer the following questions:

1. How was the sample selected for this survey?

2. Did the researchers use random or non-random sampling methods?

3. What are the advantages of their approach?

4. What are the disadvantages of this approach?

5. The sample size was 44,177. Why was the sample size so large and was this necessary?

4.4 Sample size

For the details on how to calculate the required sample size you should refer to The NIHR RDS EM / YH Resource Packs: '*Using Statistics in Research*' and '*Sampling'*. Bear in mind that if you intend to carry out a survey to calculate a proportion or a mean figure then you will need to consider the width of the confidence interval around the figure you produce. The confidence interval is the margin of error that surrounds your result. No survey can produce a result that is precisely correct. By sampling, we are aiming to achieve a degree of acceptable accuracy. There is likely to be a margin of error around any figure that you produce as a result of a survey. For example, if we survey 300 patients randomly selected from a general practitioner's records and find that 75% say that they were satisfied with the care they received at their last consultation, we can be 95% certain that the true answer is 5% either side of 75%, (i.e., between 70% and 80%). This is known as the 95% confidence interval.

Table 1 (Page 18) can be used to calculate the required sample size for your survey given any particular percentage response. For example, if we plan to repeat the above survey and know that last time we had a response rate of 75%, we find the figure of 75% in the band at the top of the table and we then move down the table column until we find an acceptable margin of error in that column and an acceptable sample size in the left-hand column.

If we decide that a margin of error of 5% either way is too great and we need a greater level of accuracy, e.g. 2%, then we move down the column headed 75% until we hit 2% (1.9% is the nearest figure). In the left-hand column we can see that we will need a sample size of 2000 to achieve this.

Remember also that you will need to allow for the expected non-response to your survey and add this to your initial sample size. For instance, if you guess that only 50% of your sample will respond to your survey, then you will need to double your required sample size.

If in doubt about what percentage figure to expect in your results, then assume the outcome to be 50%, this is represented by the far right-hand column of the table and gives the worst-case scenario for estimating sample size. So, for instance, if we want to find out what percentage of patients self treat with medicines they purchase themselves before visiting the doctor and we cannot guess the answer, we should assume it is 50% for the purpose of calculating the sample size for our survey. Assuming that we want a very precise answer which is plus or minus 2%, then using the far right-hand column of the table shows that we will require a sample size of 2500.

Use this table if you want to be 'reasonably sure' (20 to 1 odds i.e. at 95% confidence interval) of not exceeding the range of error.

Table 1

Sample Size	PERCENTAGE OF RESPONDENTS GIVING A CERTAIN ANSWER															
	1% or 99%	2% or 98%	3% or 97%	4% or 96%	5% or 95%	6% or 94%	8% or 92%	10% or 90%	12% or 88%	15% or 85%	20% or 80%	25% or 75%	30% or 70%	35% or 65%	40% or 60%	45% or 55%
RANGE OF ERROR (PERCENTAGE PLUS OR MINUS) IS SHOWN BELOW																
25	4.0	5.6	6.8	7.8	8.7	9.5	10.8	12.0	13.0	14.3	16.0	17.3	18.3	19.1	19.6	19.8
50	2.8	4.0	4.9	5.6	6.2	6.8	7.7	8.5	9.2	10.1	11.4	12.3	13.0	13.5	13.9	14.1
75	2.3	3.2	3.9	4.5	5.0	5.5	6.2	6.9	7.5	8.2	9.2	10.0	10.5	11.0	11.3	11.4
100	2.0	2.8	3.4	3.9	4.4	4.8	5.4	6.0	6.5	7.1	8.0	8.7	9.2	9.5	9.8	9.9
150	1.6	2.3	2.8	3.2	3.6	3.9	4.4	4.9	5.3	5.9	6.6	7.1	7.5	7.8	8.0	8.1
200	1.4	2.0	2.4	2.8	3.1	3.4	3.8	4.3	4.6	5.1	5.7	6.1	6.5	6.8	7.0	7.0
250	1.2	1.8	2.2	2.5	2.7	3.0	3.4	3.8	4.1	4.5	5.0	5.5	5.8	6.0	6.2	6.2
300	1.1	1.6	2.0	2.3	2.5	2.8	3.1	3.5	3.8	4.1	4.6	5.0	5.3	5.5	5.7	5.8
400	.99	1.4	1.7	2.0	2.2	2.4	2.7	3.0	3.3	3.6	4.0	4.3	4.6	4.8	4.9	5.0
500	.89	1.3	1.5	1.8	2.0	2.1	2.4	2.7	2.9	3.2	3.6	3.9	4.1	4.3	4.4	4.5
600	.81	1.1	1.4	1.6	1.8	2.0	2.2	2.5	2.7	2.9	3.3	3.6	3.8	3.8	4.0	4.1
800	.69	.98	1.2	1.4	1.5	1.7	1.9	2.1	2.3	2.5	2.8	3.0	3.2	3.3	3.4	3.5
1000	.63	.90	1.1	1.3	1.4	1.5	1.7	1.9	2.1	2.3	2.6	2.8	2.9	3.1	3.1	3.2
1200	.57	.81	.99	1.1	1.3	1.4	1.6	1.7	1.9	2.1	2.3	2.5	2.7	2.8	2.8	2.9
1500	.51	.73	.89	1.0	1.1	1.2	1.4	1.6	1.7	1.9	2.1	2.3	2.4	2.5	2.5	2.6
2000	.44	.61	.75	.86	.96	1.0	1.2	1.3	1.4	1.6	1.8	1.9	2.0	2.1	2.2	2.2
2500	.40	.56	.68	.78	.87	.95	1.1	1.2	1.3	1.4	1.6	1.7	1.8	1.9	2.0	2.0
3000	.36	.51	.62	.71	.79	.87	.99	1.1	1.2	1.3	1.5	1.6	1.7	1.7	1.8	1.8
4000	.31	.44	.54	.62	.69	.75	.86	.95	1.0	1.1	1.3	1.4	1.4	1.5	1.5	1.6
5000	.28	.40	.49	.56	.62	.68	.77	.85	.92	1.0	1.1	1.2	1.3	1.4	1.4	1.4
7500	.23	.32	.39	.45	.50	.55	.62	.69	.75	.82	.92	1.0	1.1	1.1	1.1	1.1
10000	.20	.28	.34	.39	.44	.48	.54	.60	.65	.71	.80	.87	.92	.95	.98	.99
15000	.16	.23	.28	.32	.36	.39	.44	.49	.53	.59	.66	.71	.75	.78	.80	.81
25000	.12	.18	.22	.25	.27	.30	.34	.38	.41	.45	.50	.55	.58	.56	.62	.62
50000	.08	.11	.14	.16	.17	.19	.22	.24	.26	.29	.32	.35	.39	.39	.39	.40

Now consider your own research question and aims and objectives.

1. If a survey is a possible option, think about what type of methodology you might use and why.

2. Now think about some of the key questions that you might want to ask, and write these down below.

3. A pilot study is essential in planning a survey. Write down how you would carry out a pilot study and why.

4. How would you go about selecting your participants for interview in a pilot study and why?

5. What steps might you take to increase your response rate in the main stage of your fieldwork?

5. Questionnaire design

Questionnaires are a very convenient way of collecting useful comparable data from a large number of individuals. However questionnaires can only produce valid and meaningful results if the questions are clear and precise and if they are asked consistently across all respondents. Careful consideration therefore needs to be given to the design of the questionnaire.

All questionnaires should take into account:

- *whether the questionnaire will be self-completed.* Questionnaires can be administered face-to-face by an interviewer, by the telephone or completed independently by the participants. The distinction between these methods is important because it has profound effects on the questionnaire design. A questionnaire which is to be completed by the respondent needs to be very clearly laid out with no complex filtering and simple instructions. Whereas a questionnaire to be administered by an interviewer can be much more complex;
- *the literacy level of the respondents.* Obviously respondents with low literacy levels will have greater difficulty completing a self-completion or postal questionnaire. In this case a face-to-face or telephone interview survey would be advisable;
- *the expected response rate.* The more motivated the respondent, the more likely you are to get a questionnaire returned in a postal survey. If you anticipate a very good response rate then a postal survey may do. If on the other hand, you expect a low response rate, then a personal interview survey is likely to achieve higher acquiescence;
- *the resources available.* One person would take a very long time to interview 1,000 people, however one person could carry out a postal survey of the same number of respondents with relative ease;
- *topic and population of interest.* Finally it is important to bear in mind that you do not need to reinvent the wheel. Increasingly there are ready-made questionnaires and scales available to measure patient need and outcome. Many of these are commonly used, have been well validated and can also offer useful normative data for comparison. These will be discussed in more detail later in the resource pack.

The title

All questionnaires require a title. It needs to be short, simple, appealing and inviting, not the full academic title of the study.

The identifier

Each questionnaire will also probably require a participant identifier. If your survey is confidential i.e. you know the identity of each respondent but their identity is confidential to you, then all questionnaires will require a **confidential** unique identifier. Names and addresses should not appear on the questionnaire itself. If however your survey is anonymous then you cannot know the identity of any of your respondents and none of your questionnaires should have an identifier.

Instructions

It is crucial to include instructions on the questionnaire, in terms of ticking boxes, circling numbers, allocating priority order to a list, etc. If you are doing a postal survey or a questionnaire for self-completion, you will need clear instructions for the respondent. Alternatively, if using interviewers you will need to provide them with instructions in terms of filtering and what to read out, etc.

Unless you use a ready made questionnaire, you will have to design your own. The style and content of a questionnaire will depend very much on your research question and your aims and objectives. A very **structured interview** or **postal questionnaire** will contain a greater proportion of **closed questions** with **pre-coded answers**, whereas a questionnaire or topic guide for use in a **semi-structured interview** will contain more **open-ended questions.**

Closed questions

A **closed question** is one where the possible answers are defined in advance and so the respondent is limited to one of the **pre-coded responses** given. For example, if you were to ask an asthma sufferer:

"Which of the following types of inhaler do you currently use?"

(INTERVIEWER TO READ OUT)

A dry powder inhaler	1
A metered-dose inhaler or	2
A breath-activated metered-dose inhaler	3

This is an example of a **closed question** where the possible pre-codes would be read out or shown to the respondent on a card. The choice of answers is limited to those shown on the card.

Open-ended questions

An **open-ended question** does not constrain possible responses. For example, you could ask the same patient:

"How does having asthma affect your daily life?"

In this **open-ended question** the respondent is allowed to interpret the question in their own way. They could, for instance, choose to talk about how having asthma interferes with their work, or how they can't go jogging in the winter or how they feel using an inhaler in front of school friends. The answers you get back will be very rich in details but it may be difficult to compare the responses over a large number of participants because the question is not very directed. We haven't specified to the respondents, in this case, which areas of their lives we are interested in.

A problem with asking an **open-ended** question of lots of people is that it can produce lots of different answers, which can be difficult and time consuming to code.

Partial pre-coding

This can be partially solved by using an open-ended question with **partial pre-coding**. There may be questions that we wish to ask in an open-ended manner so that we are not leading the respondents in any way, but nevertheless we can anticipate some of the possible responses. Potential responses can be anticipated by:

- carrying out a pilot study
- assessing previous studies
- guessing

The important point to note is that the pre-codes listed do not need to be exhaustive. We can always allow for an 'other, please specify' option to catch the response which we hadn't thought of in advance. An example of an open-ended question with partial pre-coding would be:

"What do you think is the cause of your asthma?"

DO NOT READ OUT (INTERVIEWER INSTRUCTIONS) CIRCLE ALL THAT APPLY

Air pollution	1
Traffic pollution/Cars/Exhaust fumes	2
Pollen	3
Dust/dust mites	4
Pets/dog/cat	5
Stress	6
Work/occupation	7
Certain foods/dairy products	8
Inherited from parents	9
Other, write in	_____

In this case we have anticipated some of the most likely answers, but we have also allowed for the respondents to give answers which we hadn't thought of in advance. So, although some **content analysis** is still required, the overall amount is still reduced. This will save time in the analysis. This type of open-ended question can be asked in an interview situation, either face-to-face or over the telephone. With self-completion questions, the respondents would be able to see the pre-codes which could influence their choice of answer.

> Definition:
> Content analysis is the systematic analysis of text or conversational transcripts to identify and group common themes.

Coding

As you can see from the above, wherever possible we need to allocate a numeric code to each possible answer. This is so that the answers can be entered into a computer for data analysis. It is generally easier to specify the codes in advance where you can anticipate the possible answers as in closed questions. Obviously with open-ended questions it may not be possible to anticipate all the possible answers. You will therefore have to code the open-ended questions as your questionnaires are returned. Some questions are asking for numerical data, for instance, 'How old are you?' or 'How many times has your social worker visited you in the last month?'. With questions like this you do not need to pre-code the answer, you simply leave a blank box to enter the exact number into.

The coding of other questions which include ordered categories, particularly those using bands or scales is straightforward. For instance, if you ask the question:

'How long ago is it since you last saw a dentist?'

Within the last month	1
More than a month ago but within the last 6 months	2
More than 6 months ago but within the last year	3
More than a year ago	4

You can see we have allocated each possible answer a numeric code. The important thing to note however is that the answers are in a particular order, thus the codes given to each answer reflect this order. The gaps between each code are not equal however. This is known as **ordinal** data.

Finally we come to **nominal** data. Nominal data is data where the codes allocated to the categories are purely nominal i.e. the numbers themselves do not mean anything numerically neither is there any sort of order implied by the numbers. The codes are there purely to act as numeric labels for each of the discrete categories.

So for example, we might ask on a questionnaire:

'Are you male or female?'

> Male   '1' or 'M'
> Female '2' or 'F'

The codes '1' and '2' do not mean anything numerically; they just signify that there are two different groups.

Definition:
Ordinal data is composed of categories, which can be placed in an order; however the gaps between each value are not necessarily of equal size.

Definition:
Nominal data, also known as categorical data is a set of unordered categories. Codes are assigned on an arbitrary basis and have no numeric meaning.

EXERCISE 7

State whether the following questions are open-ended or closed:

1. How old are you?

2. What is your name?

3. What is your occupation?

4. Would you say that traffic pollution has a direct effect on the number of people getting asthma attacks?

   (READ OUT)

   Yes, or No?

5. At home do you have a pillow with a synthetic or a feather filling or both?

6. Do you use a duvet at night?

   (READ OUT)

   Yes, or No?

7. Does this have a synthetic or a feather filling?

8. Do you own a furry pet?

9. Do you own a dog or a cat?

10. How many times a day do you use an inhaler?

(READ OUT)

Never

Once a day

Twice a day

Three times a day

Four or more times a day

Question wording

The wording of your questions is all-important since it has a direct impact on the outcome. Keep your questions simple, and use words that are easily understood by lay people. Avoid jargon, for instance, participants may misinterpret the word 'drug', use 'prescribed medicine' instead. Carry out the readability tests such as the **Fog Index** (See Appendix 1) and the **Flesch Reading Ease Score** (See Page 34).

It is important to ensure that questions are not too long and that they don't contain several questions in one sentence, which will only confuse the respondent. For example:

"Does having asthma restrict the type of work and sporting activities that you can do?"

If the answer were "Yes", what would it mean?

Avoid using ambiguous words and questions. Try not to assume anything about the respondent and avoid asking **leading questions**.

"We know that cigarette smoking can make asthma worse. How many cigarettes do you smoke a day?"

or

"Do you think the fumes from car exhausts are the main cause of asthma?"

Producing comparable data

You may want to compare your findings with previous studies. If so, it is important that you use directly comparable wording. Questionnaire design is one area where it does not always pay to be creative. Many routine questions that you wish to ask will have been refined, tested and validated by other organisations already. For example, if you want to ask a question on ethnic origin, consider including one of the questions used in the census by the Office for National Statistics (ONS) (www.statistics.gov.uk), which has been thoroughly tested.

When collecting numeric data try wherever possible to collect the data at the most detailed level you will need. For example, if you wish to record age, then do not present the respondents with age bands, rather ask them to give their exact ages. This will provide you with data at an **interval** level, rather than an **ordinal** level, which means that you can describe the data more fully. In this instance you would be able to describe the **mean** average age as well as the **median** and **standard deviation**, rather than the median alone. Interval level data also enables you to use more discriminating statistical tests.

Definition:
Interval data is measured on a scale where the distance between each point is equal.

Definition:
The mean is a measure of central tendency. It is calculated by summing all the values and dividing this by the number of cases to produce a mean average.

> Definition:
> The median is a measure of central tendency. If all the values are placed in order, the median is the midpoint or middle value.

> Definition:
> Standard deviation (SD) measures how widely spread the values of the data set are.
> If all the values are close to the mean there will be a small SD, ie close to zero.

On the other hand if you want to ask about data that may be sensitive such as income, or, if you think that the respondent will be unable to be accurate about the data, then it is preferable to ask the question with potential answers in banded groups. This will generate data at an ordinal level, which may be less useful, but avoids spurious accuracy.

For instance, when asking questions about frequency of activity such as:

"When did you last visit your GP to discuss your asthma?"

or

"How often do you use your 'blue' inhaler on an average day?"

It is important you use **pre-codes** wherever possible, because otherwise some of the answers may be so vague or wide-ranging that you will be unable to collate or compare the responses.

The question:

"When did you last visit your GP to discuss your asthma?"

could be pre-coded as follows:

In the last week	1
In the last 2 - 4 weeks	2
More than a month ago	3
More than 6 months ago	4
More than a year ago	5
Can't remember	6

Which of the following questions could be used in a survey of obesity?

If a question could not be used, state why it is unsuitable.

1. Do you weigh more or less than 10 stone?

2. Do you consider yourself to be obese?

3. Are you fat?

4. Do you consider yourself to be overweight?

5. Do you have a regular, balanced diet?

6. Do you ever take drugs to control your weight?

7. Has your GP ever prescribed medicine to control your weight?

8. What was the name of the drug prescribed, its dosage and how many times a day?

9. How many times a year do you start a diet?

10. What is your weight and height?

Question order

As a rule, it is best to move from the general to the particular when designing a questionnaire. Try to start with general fact-finding questions which are easily understood and will apply to everyone. These can act as a warm-up, getting people to ease themselves into the questionnaire. Then move onto more specific questions, which may filter people into different questions. Certain personal questions, such as age, social class and ethnicity, may be left until later, near the end of the questionnaire, when a level of rapport has been established. Likewise, embarrassing or sensitive questions may be best left until nearer the end. As with in-depth interviews, the questionnaire should 'flow'.

The order of the questions is a particularly important issue to consider when planning a self-completion or postal survey. Although respondents may choose to look ahead, you must consider what is the cumulative impact of each question on the next.

Because you cannot assume anything about the respondent, it is inevitable that the questionnaire will need to contain certain filters and instructions to either the respondent or the interviewer as to where to go next. For instance:

Q3   Are you housebound?

| Yes | 1 | If Yes, | go to Q4 |
| No | 2 | If No, | go to Q5 |

Q4   What type of home help do you receive?

Overall try to keep the question topics in a logical order. It may be best to reflect the chronological order of events if necessary, but avoid over-complicated filtering, especially in self-completion questionnaires because some respondents will be unable to follow it. If the questionnaire is self completed try to get the most important questions on the front page in case it is returned to you only partially completed.

EXERCISE 9

In which order would you place the following questions in a self-administered questionnaire in a survey of stress amongst social workers?

1	Is stress affecting your personal life?
2	In which department do you work?
3	Which is the best way of coping with stress?
4	How many home visits a week do you have?

Dealing with sensitive questions

Apart from leaving sensitive questions to near the end of the questionnaire, there are other ways of trying to elicit an honest answer.

It is possible to introduce a question by reference to the activity of others, such as:

"Not everybody uses their inhalers as often as their doctors have told them to do for a variety of reasons. Do you use your inhaler as often as your doctor has told you to?"

It may also be possible to use indirect questioning by referring to a third person before asking the respondent directly. Also, the truthfulness of responses correlates with proximity to the researcher. Face-to-face will elicit different results than web based.

Using scales in questionnaires

One way of ensuring that a question is asked in a fair and balanced way is by the use of scales. Scales will also assist in measuring the strength of attitude/feeling rather than simply "Yes" or "No", "Agree" or "Disagree".

It should be remembered however that a scale is not a precise measure of an attitude, merely a way of assessing relative measures. There are a variety of different scales to choose from.

Likert scale

The **Likert scale** is one of the most commonly used scales. Respondents are presented with one or more attitudinal statements and asked to score each statement on a multi-point scale. For instance:

"To what extent do you agree with the following statement?"

Statements	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

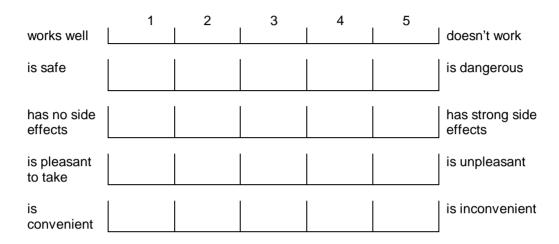|	5	4	3	2	1
Traffic pollution is a major cause of asthma					
People with asthma who smoke a lot are more likely to have worse asthma					
I believe that the medicine prescribed for me by my doctor works well					

This is an example of a five-fold Likert scale; it is also possible to have a seven-fold Likert scale. Likert scales are very popular with researchers, and sometimes they succumb to a temptation to sum the results of all the scales into a single score. This can be potentially misleading.

There is some evidence to indicate that respondents may be more likely to disagree with a negative statement than agree with a positive statement (Cohen et al, BMJ 1996; 313:841 - 4). So you need to think very carefully about how you word your statements and ensure that they are comparable with any other key studies with which you may want to compare your findings. It may also be the case that you don't need a neutral eg if one wanted to ask for opinions regarding their foster carer's support, everyone responding would lean one way or the other even if only slightly. It is easy for respondents to run their pen down neutrals without even reading the question!

Semantic differential scale

Developed by Osgood in 1957, **semantic differential scales** are used to rate individual statements, on a number of different dimensions. For example:

"Do you think that the medicine that the doctor has prescribed for your asthma:

	1	2	3	4	5	
works well						doesn't work
is safe						is dangerous
has no side effects						has strong side effects
is pleasant to take						is unpleasant
is convenient						is inconvenient

Semantic differential scales only work well when the concepts at either end of the scale are mutually exclusive. If the respondent feels that they could select both ends of the scale then the scale is impossible to answer. It is therefore very important to pilot the scales carefully.

Visual analogue scale

As an alternative to a verbal scale, a **visual analogue scale** is simply a way of asking respondents to indicate their choice visually or spatially. For instance:

"Do you think that traffic pollution has a bad effect on your asthma?"

Please place a cross on the line below.

Strongly agree	Don't know	Strongly disagree

A visual analogue scale is very difficult to interpret and is not recommended.

Ranking

A further way of getting respondents to express attitudes is by using **ranking.**

Respondents could be given a list of items on a show card and asked to rank them verbally or they could be given a number of cards each with an item on and asked to sort the cards into rank order, such as:

"Which of these items do you think has the worst effect on your asthma?"

> Exercise
> Traffic pollution
> Stress
> Diet
> Pollen

Alternatively the respondents could be asked to allocate the numbers 1 - 5 beside each category.

Indices

Sometimes discrete questions are combined to create some kind of aggregate or composite score or index. Indices are tricky things to create because it is difficult to know what level of weighting should be assigned to each discrete element of the questionnaire. Researchers often make the mistake of treating the final score as **interval data** rather than **ordinal** data.

Pilot study

Before you can embark on the main stage of fieldwork, it is crucial that the draft questionnaire is piloted. You should never use a questionnaire which has not been piloted, particularly if the questionnaire is designed for self-completion and there will be nobody around to clear up misunderstandings. A pilot stage will enable you to ensure that:

- all the relevant issues are included
- the order is correct
- ambiguous or leading questions are identified
- your pre-codes are correct, and
- you have not forgotten or omitted some issue which is really important to the respondent

The ideal situation is to test the questionnaire on a small number of respondents who are the same type as those in your sampling frame, between 5-50 respondents depending upon the final sample size. However, if the real participants are difficult to access or few in number, then you may have to test the questionnaire with slightly different participants. At the very minimum, you could try out the questionnaire on your colleagues or friends. This will at least allow you to see if the filtering and order is correct. Do remember however that if the questionnaire is destined for members of the public that you should test the questionnaire out on a lay person in preference to a professional colleague. Jargon which you use on a day to day basis may completely baffle members of the general public!

6. Using questionnaires in postal surveys

Postal surveys are a useful option to consider when selecting your method. They are usually cheaper than personal interviewing and quicker especially if the sample is large. For any postal survey regardless of the sample size you must allow at least six weeks for the first wave of questionnaires to be returned, and another four weeks for each successive mailing. As with telephone interviewing, a postal survey is useful if your respondents are widely distributed. However, due to the lack of personal contact between the respondent and the researcher, the design and layout of the questionnaire is all important. Given the popularity of postal surveys, we have included here, a very practical section on how to conduct a postal survey.

6.1 Some practical aspects of carrying out postal surveys

The **questionnaire** must be totally self-contained and self-explanatory. This means questionnaires with very complex flows and routing are not suitable. Any instructions on routing through the questionnaire, for example:

 "If the answer to Q2 is YES, go to Q6",

need to be clearly stated. Sometimes the use of coloured arrows may help this process.

Note:
The average questionnaire circulated by academics in the UK requires a reading age of 19.

A questionnaire for self-completion must look inviting and user friendly. If it appears too complex or too long, respondents simply will not bother. Obviously you need to be sure that your **sampling frame** has a sufficient literacy level before you can consider the use of a postal survey. If you think that your target population is likely to have a low literacy level, choose another method. The Questionnaire should also be in the appropriate language for that participant.

However, even with a literate population, it is necessary to cater for the lowest common denominator. Never use lengthy, complicated words when simple words will do. Be careful to avoid professional jargon and try to use a lay person's language wherever possible when addressing the general public. Test the literacy level of the questionnaire first using the **Fog Index** or the **Flesch Reading Ease Score.**

> Definition:
> The Fog Index is a calculation used to assess the required reading age of written material. (See Appendix 1 for details of how to calculate this).

> Definition:
> The Flesch Reading Ease Score computes readability based on the average number of syllables per word and the average number of words per sentence. Scores range from 0 (zero) to 100. Standard writing usually scores between 60 and 70. The higher the score, the greater the number of people who can easily understand the document.
> (The Flesch Reading Ease Score along with other readability statistics, is automatically calculated when using Microsoft Word software).

If you are doing a survey of elderly people then you may wish to consider increasing the size of the typeface used in the questionnaire.

The length of the questionnaire has to be carefully considered when conducting a postal survey but if the subject matter is of sufficient interest to the respondent, then length is less important.

It is possible to achieve high response rates with a postal questionnaire provided you adhere to the following points:

- the subject matter must be of interest to the respondent and not too sensitive
- the sampling frame must be up to date
- you carry out two waves of mailings
- you provide a pre-paid self-addressed envelope for the replies
- you reassure the respondents of their confidentiality in a covering letter
- the questionnaire is well laid out, not too complicated and is self-explanatory
- you allow sufficient time for the responses - ideally four to six weeks
- you do not mail over holiday periods such as Christmas

All mailed questionnaires should be accompanied by a covering letter. This letter will need to include:
- who you are and your role in the survey
- what the survey is about and its potential benefits
- why the respondent has been selected
- reassure the respondent about their **confidentiality** or **anonymity** (for example, you could say 'All the answers will be grouped together so that it will not be possible to identify your individual answers from the findings')
- details of where to return the questionnaire and a number to ring if they want to check the status of the survey

Definition:
An anonymous survey is one in which respondents are not allocated ID numbers and cannot be identified in any way. Sometimes very sensitive topics require anonymous rather than confidential surveys.

Definition:
A confidential survey is one in which the respondents cannot be identified except with a unique ID number used to link responses back to an individual.

Other points to consider to increase response rates include:

- use of coloured paper
- sponsorship, for example covering letter signed by somebody influential, such as a local GP
- using stamps rather than a franked envelope
- advance warning, using a letter or postcard
- using personally addressed covering letter
- using small incentives such as vouchers, stamps, pens, donations to charity. (You need to be aware that a Local Research Ethics Committee may view an incentive as an unfair inducement to participate, so your incentive should not be too generous)

6.2 Postal Survey Reminders
If two waves of mailing are proposed, it is necessary to make an estimate of the likely response rate to the first wave when deciding on the total number of questionnaires and envelopes required. Use the table below to guide your decision and remember that:

- every questionnaire will require two envelopes, one to be sent out and one for the reply

- if you know who hasn't responded to the first mailing they will need to be sent a reminder letter plus a questionnaire plus a reply envelope
- if the questionnaires are anonymous then you must send a second mail out

For instance, if we mail 100 respondents and get responses initially from 40%, we will need to mail reminders to 60 people. So, in total we would require 160 questionnaires to achieve a response rate of 50%.

Table 2

If your initial response rate is:	You will need this many questionnaires per 100 sample members	You will need this many envelopes per 100 sample members
50%	160	320
60%	150	300
70%	140	280
80%	130	260
90%	120	240

However, if our initial response rate is higher, e.g. 80%, we would only need to send reminders to 20 people, to achieve a response rate of 90%. If we can estimate the response rate in advance we can minimise the number of questionnaires and envelopes required. A pilot study or an assessment of previous studies may help you to decide on your anticipated response rate.

7. Data analysis

Finally we need to consider the analysis of the data. It is all too easy to busy oneself with the questionnaire design and data collection and not give any thought to the analysis stage. However it is actually crucial to think about the analysis very early on. It is important to establish early on the exact data type of the main outcome measures, e.g. nominal, ordinal, or interval. The data type in turn determines which type of statistical test is most appropriate, and this in turn has implications for the required sample size.

Once the data has been assigned a numerical code as discussed earlier, it needs to be entered into the computer software that you will be using for data analysis. This is likely to be SPSS (See The NIHR RDS EM / YH Resource Pack: '*Using SPSS*'). Once The data has been entered it is necessary to check the data for errors and typos. Although it is possible to carry out some simple hand counts and therefore avoid computers, this approach is really not recommended. If your survey consists of more than 20 cases and you want to cross-tabulate the answers to two or more questions it becomes practically impossible to do this by hand.

Once you are satisfied that the data you have entered onto the PC is correct, you will probably want to start your analysis by requesting basic frequencies. These are simple counts of the number of individuals answering each question. The next stage is to request **cross-tabulations** or **contingency tables** which enable you to analyse the way that answers to one question might vary with answers to another question. For example:

Gender by consultation with GP in last year

	Male	Female	Total No.
Consulted GP in last year	52	87	139
Not consulted GP in last year	57	13	70
Total No.	109	100	209

We can see in the table above that there would be a difference in the consulting behaviour between men and women. However, just by looking at this we cannot tell if this difference is statistically significant. The next stage is to apply some statistical test to the data to see if there is a statistically significant difference between men and women in our sample in terms of consultations. In this particular example, the appropriate test to use would be a chi-square test because both variables are nominal. It is important to understand that the most appropriate statistical test is likely to depend upon the type of data you are collecting and the number of groups that you are comparing. It is not possible to cover the statistical tests in detail here and you are advised to consult The NIHR RDS EM / YH Resource Pack '*Using Statistics in Research*' for further details.

The important point to remember is that you do not need to know how to calculate these tests by hand because the computer will do this for you. However, you do need to know which is the appropriate test to request and how to interpret the answer.

Now consider your own research question and aims and objectives.

1) If a survey is a possible option, think about what type of methodology you might use and why.

2) Now think about some of the key questions that you might want to ask, and write these down below.

3) A pilot study is essential in planning a survey. Write down how you would carry out a pilot study and why.

4) How would you go about selecting your subjects for interview in a pilot study and why?

5) What steps might you take to increase your response rate in the main stage of your fieldwork?

Summary

In this resource pack we have concentrated on five main areas:

1. Why surveys are useful.

2. Different methods of how to carry out a survey.

3. How to use questionnaires.

4. Aspects of questionnaire design, and

5. Using questionnaires in a postal survey.

You should by now be able to describe the advantages and disadvantages of the following methods:

- face-to-face interviews
- telephone interviews
- postal surveys, and
- self-completion questionnaires including ready-made instruments

You should be able to list the factors which may increase participation and/or response rate.

You should understand the difference between open-ended and closed questions and know what a pre-code is.

You should also be able to distinguish between a structured and a semi-structured questionnaire. As you will recall, a structured questionnaire with a majority of closed questions and pre-coded answers is appropriate when you trying to directly compare the responses of a large number of people. Whereas a semi-structured questionnaire will allow you to ask more open-ended questions which are rich in detail but more difficult to analyse and compare.

Remember when designing your questionnaire to avoid:

- Using ambiguous or multiple questions
- Using double negatives
- Using long, complicated sentences or jargon
- Using leading questions

Finally you should be able to list the factors which may increase participation and/or response rate.

Answers to exercises

Exercise 1
Possible answers are:

Main Advantages
- the ability to carry out a longer interview on a more complex topic, with complicated filtering
- a higher response/completion rate
- the ability to use visual aids and prompts
- the opportunity to develop a rapport
- can reach some people that due to disability or socio/economic status may not normally be reached

Main Disadvantages
- labour intensive for large samples
- time consuming
- difficult to contact people at home
- difficult to get honest answers to personal/sensitive questions such as drug-taking, diet, etc

Exercise 2
Possible answers are:

Main Advantages
- can build rapport
- easier to contact geographically-spread respondents
- may be easier to contact professionals
- can be cheaper and quicker than a face-to-face survey
- high completion rate if person commits to questionnaire

Main Disadvantages
- not appropriate if sampling frame do not all have equal access to a telephone
- sampling frame - even those with a telephone may be ex-directory
- cannot use visual aids
- interview may be impersonal
- interview length is limited

Exercise 3
Possible answers are:

Main Advantages
- not labour intensive, possible to mail a large sample, and therefore may be cheaper than face-to-face or telephone
- gives people the opportunity to answer questions in their own time; they may need to refer to documents etc.
- useful for a widely distributed sample

Main Disadvantages
- possibly lower response depending on topic and motivation
- usually takes much longer, need to give people sufficient time to reply and send out mailers
- not suitable for people with literacy problems

- cannot guarantee that the person who completed it was the target participant

Exercise 4
1. Stratified random sample. The sample is stratified because the sample has been selected to ensure that two different groups are represented.
2. Disproportionate stratified random sample. This sample is stratified to ensure that patients from the two different groups are picked up, however the two groups are selected so that they are equal in size and are not representative of the patient base.
3. Quota. The sample is not randomly selected but the respondents are selected to meet certain criteria.
4. Convenience. The sample is not randomly selected and no quotas are applied.
5. Systematic sample.
6. Cluster sample. The patients are selected only from certain wards.

Exercise 5
1. The researchers used a convenience sampling approach, i.e. they selected people on the basis that they were easy to access. Respondents were therefore self-selected.
2. The sampling method used was non-random.
3. The advantage of this approach was that they were able to obtain the views of a large number of people very quickly and easily with little expense.
4. Unfortunately the convenience sample approach means that the sample is not representative of the population of individuals with asthma. Because a large part of the survey is made up of people attending surgeries or pharmacies, the sample will tend to over-represent those individuals requiring the most treatment. It will also over-represent those individuals who are most interested in expressing their opinions.
5. The sample achieved was very large because it was self-selected, and therefore the researchers would have had little control over how many people participated.

Exercise 6
Since this exercise is based on a self selected example it is not possible to provide specific answers to the questions. You should, however, draw on the guidelines in the text when approaching them.

Exercise 7
1. Open
2. Open
3. Open
4. Closed
5. Closed
6. Closed.
7. Closed.
8. Closed, but will probably attract some strange answers and questions!
9. Closed.
10. Closed.

Exercise 8
Q1 does not provide all possible alternatives. What happens if the respondent weighs exactly 10 stone?
Q2 The term 'obese' is jargon and not all respondents would understand its true meaning. (One lady who was told that she was obese by her GP thought he told her that she was a 'beast'). A preferable term would be over-weight.
Q3 The term 'fat' is derogatory and should not be used at all.
Q4 is acceptable.
Q5 is ambitious and vague. What is meant by regular? What is meant by balanced? The question is too subjective.

Q6 is open to misinterpretation. The term 'drug' has a different meaning in lay use compared to its clinical definition.
Q7 is acceptable.
Q8 is three questions in one
Q9 is a leading question and assumes that the respondent uses diets.
Q10 asks 2 questions in one. It would be better to separate them.

Exercise 9
The correct order would be:
2, 4, 1, 3

Exercise 10
Since this exercise is based on a self selected example it is not possible to provide specific answers to the questions. You should, however, draw on the guidelines in the text when approaching them.

Further reading and resources

Recommended Reading

Boynton PM, Greenhalgh T. Selecting, designing and developing your Questionnaire. BMJ, 328:1312-15, 2004.

Boynton PM. Administering, analysing and reporting your Questionnaire. BMJ, 328:1372-75, 2004.

Boynton PM, Wood GW, Greenhalgh T. Reaching beyond the white middle classes. BMJ, 328:1433-36, 2004.

Bradburn. Asking Questions: The definitive guide to Questionnaire design. Wiley, 2004.

Crombie I, Davies H. Research in Health Care. Chichester: John Wiley & Sons Ltd, 1996.

Dillman DA. Mail and Internet Surveys: The Tailored Design Method. Wiley & Sons, 2006.

Drury M, Carter Y, Thomas C. Research Methods in Primary Care. Oxon: Radcliffe, 1996.

Foddy W Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research Cambridge University Press, 1993.

Fowler FJ. Improving Survey Questions: Design and Education. Sage Publications, 1995.

Gillham WEC. Developing a Questionnaire. Continuum, 2000.

Kline P. The Handbook of Psychological Testing. London Routledge, 1993.

Moser C and Kalton G. Survey Methods in Social Investigation. Hants: Gower, 1990.

Oppenheim A. Questionnaire Design, Interviewing and Attitude Measurement. London: Pinter Publishers Ltd, 2000.

Peterson R A Constructing Effective Questionnaires Sage Publications 2000.

Sapsford R, Jupp V. Data Collection and Analysis. London: Sage Publications, 2006.

Stone D. How to Design a Questionnaire. BMJ, 307:1264-6, 1993.

Streiner & Norman Health Measurement Scales: A Practical Guide to their Development and Use Oxford University Press 1989.

Willis G. Cognitive Interviewing: A tool for improving Questionnaire design. Sage, 2004.

Examples of existing health outcome measures or indices

Please refer to the question bank at:

http://qb.soc.surrey.ac.uk/docs/surveys.htm

for examples of Social and Health care validated Questionnaires.

Glossary

Anonymity   is the protection of the identity of research participants such that even the researcher cannot identify the respondent to a questionnaire. Questionnaires in an anonymous survey do not have an identification number and cannot be linked back to an individual. Anonymity should not be confused with confidentiality where individuals can be identified by the researcher.

Association   between two variables represents some sort of relationship. An association can be a causal one or it might be spurious. Associations can be positive or negative.

Bias   is a deviation of the results from the truth. This can either be due to random error or, more likely, due to systematic error. The latter could be caused by, for example, sampling or poor questionnaire design.

Case   is a unit of analysis. Usually this takes the form of an individual participant but it could be a different unit of analysis altogether such as a family or a blood culture.

Categorical data   see nominal data.

Closed question   is one where the possible answers have been defined in advance and so the respondents' answers will be restricted to pre-coded responses offered. A pilot study should be carried out to decide on the correct pre-codes.

Coding   is the process by which responses to questionnaires or other data is assigned a numerical value or code in order that the data can be transferred to a computer for data analysis. See also 'pre-codes', 'closed questions', 'open-ended questions' and 're-coding'.

Cohort design   is a longitudinal design where the same individuals are interviewed or observed repeatedly over time. Respondents usually share a common characteristic.

Confidentiality   is the protection of the identity of research participants so that identities cannot be revealed in the research findings and the only person who can link a respondent's completed questionnaire to a name and address is the researcher. A questionnaire with just a coded identification number is confidential. This should not be confused with anonymity, where not even the researcher can identify the participants.

Confounding variable   is one which systematically varies with the independent variable and also has causal effect on the dependent variable. The influence of a confounding variable may be difficult to identify, since it is sometimes difficult to separate out the independent variable from any confounding variables in real life.

Constant error   can be caused by the presence of a confounding variable in an experiment. It is also an alternative term for systematic bias.

Construct validity   is the extent to which the measurement corresponds to the theoretical concepts (constructs) concerning the object of the study.

Content validity   is a set of operations or measures which together operationalize all aspects of a concept.

Criterion validity   is the extent to which measurement correlates with an external indicator of the phenomenon. There are two types of criterion validity - concurrent and predictive: i) concurrent validity is a comparison against another external measurement at the same point in time ii) predictive validity is the extent to which the measurement can act as a predictor of the criterion. Predictive validity can be useful in relation to health since it can act as an early risk indicator before a condition develops in full.

Cross-sectional design   is analogous to a snap-shot.   A cross-sectional design is one which focuses on a single fixed period in time, and can provide a description of respondents that differ on a number of variables.

Dependent variable   is known as the outcome variable. The value of a dependent variable is dependent on other independent variables and its value will change as the independent variable or intervention changes. Statistical techniques can be used to predict the value of the dependent variable. An example of a dependent variable might be blood pressure.

Descriptive design   is one which seeks to describe the distribution of variables for a particular topic. Descriptive studies can be quantitative, for instance, a survey, but they do not involve the use of a deliberate intervention.   However, it is possible to carry out correlational analysis of the existing variables in a descriptive study.

Descriptive statistics   are used to describe and summarise variables within a data set including describing relationships between variables.   They do not seek to generalise the findings from the sample to the wider population, unlike inferential statistics.

Error   can be due to two sources: random error and systematic error. Random error is due to chance, whilst systematic error is due to an identifiable source such as sampling bias or response bias.

Experimental design   is one in which there is direct control over the use of an intervention. In a classic experimental design, the participants are randomised into intervention and control group and the dependent variable is assessed before and after treatment. See 'RCT'.

External validity   relates to the extent to which the findings from a study can be generalised (from the sample) to a wider population (and be claimed to be representative).

Extraneous variable   is a variable other than the independent variable which may have some influence on the dependent variable and may be a potential confounding variable if it is not controlled for.

Face validity   is the extent that the measure or instrument being used appears to measure what it is supposed to. For example, a thermometer might be said to possess face validity.

Hawthorne Effect   is the change that occurs in a participant's behaviour or attitude as a result of being included in the study and being placed under observation. The term derives from industrial, psychological studies which were carried out at the Hawthorne plant of the Western Electric Corporation in Illinois in the 1920s and were reported by Mayo. He found that whatever experimental environmental conditions were tried out on the workers, productivity always went up. He realised that it was the effect of actually being under study that resulted in a change of behaviour and so increased productivity.

Hypothesis   is a statement about the relationship between the dependent and the independent variables to be studied. Traditionally the null hypothesis is assumed to be correct, until research demonstrates that the null hypothesis is incorrect. See 'null hypothesis'.

Incidence   can be defined as the number of new spells of a phenomenon e.g. illness, in a defined population in a specified period. An incidence rate would be the rate at which new cases of the phenomena occur in a given population. See 'prevalence'.

Independent variable   is one which 'causes' the dependent variable. The independent variable takes the form of the intervention or treatment in an experiment and is manipulated to demonstrate change in the dependent variable.

In-depth interview   takes an unstructured, qualitative approach. The questions asked will be mostly open-ended and overall the degree of control over both the order and content of the interview is less than in a structured interview.

Indexing   is a process of collating indicators to create a single index of a particular phenomenon such as mental health, quality of life, daily functioning, etc.

Indicator   is the operationalized form of a concept. In research, concepts need to be tightly defined so that they can be measured. To measure a concept we have to translate it into a specific indicator.

Instrument validity   is the extent to which the instrument or indicator measures what it purports to measure. Note that a study could have instrument validity but still lack validity overall due to lack of external validity.

Internal validity   relates to the validity of the study itself, including both the design and the instruments used.

Interval data   is measured on an interval scale where the distance between each value is equal and the distance between values is the same anywhere on the scale. Interval level data does not possess a true zero. Unlike ratio data.

Intervening variable   occurs in the causal pathway between the independent variable and the dependent variable. It is statistically associated with both the independent and the dependent variable.

Longitudinal study   is one in which groups of people are interviewed repeatedly over a period of time. Where the same group of people are followed up over time this is known as a cohort study. However, if a group of different people are interviewed in each wave of a survey this is known as a trend design.

Mean   is a measure of central tendency. It is calculated by summing all the individual values and dividing this figure by the total number of individual cases to produce a mean average. It is a descriptive statistic which can only be applied to interval data.

Median   is a measure of central tendency. It is the mid-point or middle value where all the values are placed in order. It is less susceptible to distortion by extreme values than the mean, and is a suitable descriptive statistic for both ordinal and interval data.

Mode   is a measure of central tendency. It is the most frequently occurring or most common value in a set of observations. It can be used for any measurement level but is most suited for describing nominal or categorical data.

Nominal data   also known as categorical data, is a set of unordered categories. Each category is represented by a different numerical code but the codes or numbers are allocated on an arbitrary basis and have no numerical meaning. See also 'ordinal' and 'interval data'.

Null hypothesis   is the alternative hypothesis. It usually assumes that there is no relationship between the dependent and independent variables. The null hypothesis is assumed to be correct, until research demonstrates that it is incorrect. This process is known as falsification.

Open-ended question   is one which allows the respondent the freedom to give their own answer to a question, rather than forcing them to select one from a limited choice. Open-ended questions are commonly used in in-depth interviews, but they can also be used in quantitative structured interviews as well.

Ordinal data   is composed of a set of categories which can be placed in an order. Each category is represented by a numeric code which in turn represents the same order as the data. However, the numbers do not represent the distance between each category. For instance, a variable describing patient satisfaction may be coded as follows: Dissatisfied 1, Neither 2, Satisfied 3. The code 2 cannot be interpreted as being twice that of code 1.

Prevalence   is the number of cases of participants with a given condition or disease within a specified time period. The prevalence of a condition would include all those people with the condition even if the condition started prior to the start of the specified time period. See Incidence.

Prospective study   is one that is planned from the beginning and takes a forward looking approach. Participants are followed over time and interventions can be introduced as appropriate.

Qualitative research   deals with the human experience and is based on analysis of words rather than numbers. Qualitative research methods seek to explore rich information usually collected from fairly small samples and includes methods such as in-depth interviews, focus groups, action research and ethnographic studies.

Quantitative research   is essentially concerned with numerical measurement and numerical data. All experimental research is based on a quantitative approach. Quantitative research tends to be based on larger sample sizes in order to produce results which can be generalised to a wider population.

Questionnaire   is a set of questions used to collect data. Questionnaires can be administered face-to-face by an interviewer, over a telephone, on the web or by self-completion. Questionnaires can include closed and open-ended questions.

Random error   is non-systematic bias which can negate the influence of the independent variable. Reliability is affected by random error.

Randomised Control Trial (RCT)   is seen as the 'gold standard' of experimental design. As the name implies, participants are randomly allocated to either the intervention or the control group.

Ratio level data   is similar to interval data in that there is an equal distance between each value, except that ratio data does possess a true zero. An example of ratio data would be age.

Re-coding   is the process of altering the codes assigned to a particular variable, usually by aggregating categories. For instance, continuous interval data such as age may be re-coded into age bands, thus making it ordinal data. Re-coding allows data to be analysed and compared in different ways than in its original state.

Reliability   is concerned with the extent to which a measure gives consistent results. It is also a pre-condition for validity.

Representativeness   is the extent to which a sample of participants is representative of the wider population. If a sample is not representative, then the findings may not be generalisable.

Response rate   is the proportion of people who have participated in a study or completed a questionnaire. It is calculated by dividing the total number of people who have participated by those who were approached or asked to participate.

Retrospective design   is one which looks backwards over time, often using data already collected by others. It usually takes the form of correlational research identifying relationships between independent and dependent variables.

Sample   is a group or subset of the chosen population. A sample can be selected by random or non-random methods. Findings from a representative sample can be generalised to the wider population.

Sampling frame   is the pool of potential participants which share a similar criteria for entry in to a study. The sampling frame is also known as the 'population'.

SPSS   (Statistical Package for the Social Sciences) is an increasingly popular and easy-to-use software package for data analysis.

Standard Deviation (SD)   is a measure of dispersion, as it calculates the amount of deviation from the mean. It reflects the degree to which the values in a distribution differ from the mean.

Survey   is a method of collecting large scale quantitative data but does not use an experimental design. With a survey there is no control over who receives the intervention or when. Instead a survey design can examine the real world and describe existing relationships. A survey can be either simply descriptive or correlation's.

Validity   is the extent to which a study measures what it purports to measure. There are many different types of validity.

Variable   is an operationalized concept. A variable is a phenomenon that varies and must be measurable. An outcome variable is known as the dependent variable and the effect variable is known as the independent variable. The independent variable has a causal effect on the dependent variable.

Weighting   is a correction factor which is applied to data in the analysis phase to make the sample representative. For instance, if a disproportionate stratified sampling technique has been used then the total data may need to be re-weighted to make it representative of the total population. Weighting is also used to correct for non-response, when the respondents are known to be biased in a systematic way.

Appendix 1

The Fog Index

The Fog Index was developed by an American, Robert Gunning as a way of quickly checking the readability of text. The Fog Index aims to show the number of years of education which the reader needs to have had to understand the text. So text with an index of 12 should be clear to a reader who left school at 17 (assuming he started at age 5). Remember that the Fog Index is just a guide.

It is calculated in the following way:

- Take a piece of text of about 100 words.

- Divide the number of words by the number of sentences. This gives the average number of words per sentence.

- Count up the number of words of three syllables or more in the passage, e.g. 'ibuprofen', 'infarction', 'leukaemia'.

    Don't count words that are:
    Capitalised
    Combinations of short words
    Made into three syllables by the addition of -ed or -es e.g. admitted

- This gives you the percentage of 'hard' words in the passage.

- Add the average number of words per sentence to the percentage of 'hard' words in the passage.

- Multiply by 0.4 to get the Fog Index.