Chapter 4
Planning the research

4.1 Introduction

After deciding on the research topic, the investigators have to think carefully about the plan of the research. In this process, they consider the options they have about different ways in which the research topic can be investigated, i.e. a research design. In making this choice, they have to weigh two factors. They should try to choose a design that will give most definitive answers about the research topic. But they have to weigh this against the feasibility of doing the study. They have to consider, among other things, their own capabilities, the availability of material or subjects for the research, and the availability of resources. Often, a trade-off has to be made between the ideal and the possible. The best should not be made the enemy of the good.

After deciding on a research design that is appropriate to deal with the research topic and that is feasible, they have to look again at the broad research topic, and define and refine it into a research question which can be answered by the research design. For many studies, this will involve generating a research hypothesis that can be tested.

Among the issues the investigators have to deal with in designing the research is the question of sampling. Since the study cannot include all the target population, they have to depend on the accessible population, and select a sample that is as representative as possible of this population. The size of the sample is an important decision to make. If, on the one hand, the sample is too small, the results obtained will not be reliable, the resources for the research will be wasted and, if human subjects are involved, it would have been unethical to subject them to research that does not give useful results. If, on the other hand, the sample is too large, it prolongs the study and makes it more expensive, with no added scientific value. The investigators also have to give attention to how the study results will be measured, by choosing methods that are reliable and valid.

The design of qualitative research needs different approaches from that of quantitative research. These approaches include observation, in-depth interviews and focus group discussions. If a questionnaire is used to collect information from respondents, there are a number of options for the investigators, and there are guidelines to follow.

Last but not least, planning is the time to think carefully about ethical implications before the study is implemented.
All these topics will be discussed in the next sections. For more detail, the references and additional sources listed for the chapter can be consulted.

4.2 Types of research design

The study type may dictate certain research designs. More commonly, the study objectives can be achieved through a number of alternative designs. The investigators have to select the most appropriate and most feasible design.

Generally, there are two main categories of research design: observational study, and experimental or intervention study. In the observational study, the investigators stand apart from events taking place in the study. They simply observe and record. In the experimental or intervention study, the investigators introduce an intervention and observe the events which take place in the study.

Observational studies

An observational study may be descriptive or analytical. A descriptive study is an observational study that simply describes the distribution of a characteristic. An analytical study is an observational study that describes associations and analyses them for possible cause and effect.

An observational study may be cross-sectional or longitudinal. In a cross-sectional study, measurements are made on a single occasion. In a longitudinal study, measurements are made over a period of time.

A longitudinal observational study may be retrospective or prospective. In a retrospective study, the investigators study present and past events. In a longitudinal prospective study, the investigators follow subjects for future events.

Case–control studies are a type of observational-analytical-retrospective studies over time in which a group of subjects with a specified outcome (cases) and a group without that outcome (controls) are identified. Investigators then compare the extent to which each subject was previously exposed to the variable of interest, such as a risk factor, a treatment or an intervention. Case–control studies are useful for studying rare conditions and conditions with long intervals between exposure and outcome such as, for example, risk of developing neoplasia. In such situations, a prospective study will be difficult. Case–control studies can be efficient and economical, but do not have the strength of evidence of a prospective study.

In clinical and epidemiological research, a longitudinal observational study is usually called a cohort study. The word cohort was the ancient Roman term for a group of soldiers who marched together into battle. The prospective cohort design is generally considered to be the “crème de la crème” of observational methodologies for the following reasons.
Planning the research

- Data are gathered prospectively.
- Recall bias is not a problem (research subjects are not asked to recall past events).
- Time–order relationships are clear (it is easy to decide that an outcome followed, rather than preceded, a possible cause).
- Investigators have much more control on the quality of the data.

There are, however, some drawbacks.

- The biggest single problem of these follow-up design investigations is the loss of valuable information through attrition, due to loss to follow-up, or subjects opting out of the study.
- Subjects may change their behaviour over time.
- A bias can occur if there is unequal surveillance of subjects in the two compared groups, during follow-up.

One of the best examples of a prospective cohort study was initiated by Austin Bradford Hill and Richard Doll, to investigate the relationship between smoking and lung cancer. They followed up 40 000 British doctors who were divided into four cohorts: non-smokers, and light, moderate and heavy smokers. Death was the outcome they recorded. They used both all cause death (any death) and cause specific death (death from a particular disease). Publication of their interim 10 year results in 1964, showed a substantial excess in both mortality from lung cancer and all cause mortality in smokers, with a “dose-response” relation (that is, the more the subjects smoked the greater were their chances of getting lung cancer). The study went a long way in demonstrating that the link between smoking and ill-health was causal rather than coincidental. The 20 year and 40 year results of this momentous study (which achieved 94% follow-up of those recruited in 1951 and not known to have died) illustrate the strength of evidence that can be obtained from a properly conducted cohort study (Doll and Hill, 1964; Doll and Peto, 1976; Doll et al., 1994).

Experimental or intervention studies

In the experimental or intervention study, the investigators test the effect of an intervention on the events taking place in the study. An experimental or intervention study may be controlled or non-controlled. Giving a treatment to a patient or group of patients and finding that the treatment works gives only preliminary and non-definitive information. We do not know what would have happened if no treatment or a different treatment was given. For a more definitive answer, we need a “control” group of patients who do not get the treatment under study.
Hawthorne effect: In the late 1920s, a group of researchers at the Western Electric Hawthorne Works in Chicago were investigating the effects of lighting, heating and other physical conditions upon the productivity of workers. Much to the surprise of the researchers, the productivity of the workers kept improving even when the actual physical conditions were not improved. The Hawthorne effect can be manifested in clinical research settings. Even “inert” treatments might result in significant improvements in the patient’s condition (Polgar and Thomas, 2000).

A controlled experimental study may be randomized or non-randomized. In testing the outcome in a group of patients who receive the treatment and another group who do not, we are still not sure whether any difference observed is because of the treatment or because the characteristics of the patients in the two groups were different. The best way to be sure is to randomize the allocation of patients to either treatment or to no treatment.

Randomized controlled trials are intervention studies characterized by the prospective assignment of subjects, through a random method, into an experimental group and a control group. In a clinical trial, the experimental group receives the drug or treatment to be evaluated, while the control group receives a placebo, no treatment, or the standard of care. Both groups are followed for the outcome(s) of interest. Randomization is the most reliable method to ensure that the participants in both groups are similar as far as possible with respect to all known or unknown factors that might affect the outcome. With randomization, only chance determines the assignment of subjects to study groups. Random allocation does not mean haphazard allocation. It is a carefully planned method of assigning subjects to similar groups. If important risk factors can be identified at the outset, subjects may be grouped or stratified prior to assignment. Whenever it is ethical and practical, a randomized design should be considered in controlled intervention studies.

Controlled trials without randomization are intervention studies in which allocation to either experimental or control group is not based on randomization, making assignment subject to possible biases that may influence study results.

A crossover study is a special design of controlled intervention study that is sometimes used in drug trials. In this design, half of the participants are randomly assigned to start with the placebo and then switch to active treatment, while the other half does the opposite. It has the advantage of reducing the number of subjects required, since each subject serves as both an experimental subject and a control. It also decreases the biological variability inherent in comparing different subjects by comparing each subject with himself or herself. It has the disadvantage of increasing the duration of the study. There will also be a problem if the treatment has a carry-over effect after it is stopped.
A before-and-after study is a method of control in which results from experimental subjects are compared with outcomes from patients treated before the new intervention was available. These are called historic controls.

A randomized controlled trial may be blinded if participants in the trial are likely to change their behaviour in a systematic way that may influence the outcome of the study when they are aware of which intervention they receive. (Ophthalmologists prefer the term “masking” to the term “blinding”.)

Blinding can take place at a number of levels. At one level, those responsible for assigning the subjects to groups do not know to which group the next subject will be assigned. In another level, research subjects are also not aware of which intervention they are receiving. Then, health workers who take care of patients in the study may not be allowed to know what treatment the different patients are receiving. Lastly, researchers who assess the outcome are also not able to distinguish the subjects in the different groups.

The term double-blind is used when neither researchers nor subjects are aware of the type of intervention. A trial in which there is no attempt at blinding may be called open or open label.

The Rosenthal effect: Rosenthal and his colleagues in 1976 performed an experiment involving the training of two groups of rats in a maze learning task. A bright strain and a dull strain of rats especially bred for the purpose were trained by undergraduate student experimenters to negotiate the maze. After a suitable training interval, the relative performances of the groups were compared. Not surprisingly, the bright strain significantly outperformed the dull strain. What was surprising, however, was that the two strains were actually not different. The two groups of rats were actually genetically identical. The researchers had deceived the student experimenters for the purposes of the study, and the students’ expectations of the rats had resulted in different methods of treatment, which had affected the rats’ learning ability. These results have been confirmed time and time again in a variety of experimental settings, and with a variety of subjects. They confirm the need for blinding (Polgar and Thomas, 2000).

4.3 Selecting a research design

A research question may be answered by more than one research design. The researcher has to select the appropriate design for the particular study. All types of research design have a place, and all have advantages and disadvantages. But not all types of design are always possible for a particular study.

For example, the investigators may want to study if there is a relationship between post-menopausal hormone replacement therapy and subsequent development of uterine
endometrial carcinoma. The investigators can design an observational study or an experimental study. If the decision was for an observational study, the investigators may do a descriptive study or an analytical study.

For a descriptive study, they will review the clinical records of all patients diagnosed as having endometrial carcinoma. They will look for a history of post-menopausal hormonal therapy. This study will be useful but cannot be definitive. It shows whether further study is needed to confirm or refute the impression gained from the descriptive study. The information about the strength of the association will also help in the design of further analytical studies. The finding that many of the women who developed endometrial carcinoma had a history of hormonal therapy cannot lead to any conclusion. It may simply mean that this therapy is widely used in the community, both by women who develop and who do not develop endometrial carcinoma. This shows the need for further studies.

For an analytical study, the investigators may do a cross-sectional study or a longitudinal study. In a cross-sectional study, the investigators may study all post-menopausal women admitted to hospital over a defined time period. For each woman, they record whether she received or did not receive hormonal therapy, and whether she had or did not have endometrial cancer. The advantage of this study is that it can be done rapidly. It gives more evidence than the simple descriptive study. However, the two groups of patients may not be comparable.

In a longitudinal observational study, the investigators may do a prospective study or a retrospective study. For a prospective study, a cohort of two groups of post-menopausal women is followed up: one group already receiving hormone replacement therapy and another matched group not receiving this therapy. For a retrospective study, a case–control design can be selected. A group of women who have recently developed endometrial cancer (cases) and a group of women with similar characteristics and did not develop endometrial cancer (controls) are identified. The use of hormone replacement therapy in each woman in the case group and in the control group is determined to assess exposure history. The advantage is that the study can be done relatively quickly. The disadvantage is that the two groups may still not be completely similar. Other variables may influence the outcome and may be difficult to exclude.

If the investigators decide on an experimental or intervention study, they may select a randomized or a non-randomized design. In a randomized controlled study, post-menopausal women identified from a population are randomly assigned either to a study group that will receive hormone replacement therapy or to a control group that will be prescribed a placebo. Both groups will then be followed prospectively to determine how many in each group will develop endometrial cancer. This study, if successfully conducted, will provide a more definitive answer to the research question. However, it will raise ethical concerns. Additional difficulties are the large sample size needed
because of the relatively low incidence of the disease, the long follow-up because of the long latent period before the development of the disease and the possibility of poor compliance or loss to follow-up. Alternatively, a non-randomized controlled design may be considered. This may be easier, will allow women to make an informed choice but there will be a need to consider other possible variables that may influence the outcome, since the two groups may not be similar.

Different types of research design are not considered equal in the strength of evidence they provide. In the traditional hierarchy of evidence, randomized controlled studies are generally ranked high, followed by cohort and case–control studies, while observational descriptive studies are ranked at a lower level. The investigators may, however, not be able to select the design that gives a high level of evidence, because it will not be feasible to do, or will not be ethical to do. In this case, their selection of another design will be acceptable and justified.

### 4.4 Defining and refining the research question

In order to develop the research design, the research topic often has to be changed to a research question, and the research question should be defined and refined so that it can be answered with precision.

If we take again the example of the relationship between post-menopausal hormone replacement therapy and subsequent development of endometrial carcinoma, the research question will be: Does post-menopausal hormone replacement therapy predispose women to develop endometrial cancer?

For the purpose of the research design, the question needs to be better defined. The hormone replacement therapy should be specifically stated. Is it oestrogen alone or oestrogen in combination with a progestagen? Does the duration of therapy need to be defined as, for example, more than one year? Should the diagnosis of endometrial cancer be specified as histologically confirmed?

For the purpose of the research design, the question also needs to be refined. The research will only be able to determine if there is an association or not. The refined question should therefore be: Is post-menopausal hormone replacement therapy, as defined, associated with a subsequent increased risk of endometrial cancer? The association, if found, will need an explanation, but cannot be taken as meaning causation without further questioning.

If we take another example for a research question, “Is passive smoking harmful to the foetus?” the question needs to be better defined and also refined.

The first definition is about passive smoking. What arbitrary definition should be accepted, in terms of number of cigarettes smoked every day? This is called an
operational definition. The operational definition is a statement of how the researchers in a particular study choose to measure the variable in question. It should be unambiguous and have only one possible interpretation. Another definition that needs to be made is about effect on the foetus. Could it be defined as effect on intrauterine growth retardation, biophysical profile as determined by ultrasound examination, low birth weight, or the condition at birth (Apgar score for example)? Choice of any of these outcomes will affect the size of the sample to be studied. It will also need control for other variables, which will have to be excluded.

After considering these definitions, there is a need to refine the research question to be, for example, “Are the children born to women whose husbands smoke more than 20 cigarettes a day, of lower birth weight than children born to women whose husbands do not smoke”? This research question is now suitable to turn into a specific hypothesis that can provide a good basis for the development of an appropriate design and calculation of the sample size needed.

4.5 Generating the research hypothesis

If the research question is concerned with relationships between observations or variables, a research hypothesis will need to be developed. The research hypothesis is a tentative statement that can be tested by a scientific research design. Using the previous two examples, the research hypotheses could be as follows.

- Post-menopausal women who received hormone replacement therapy, of a specified type and duration, are more likely to develop endometrial cancer than post-menopausal women who did not receive such therapy.
- Children born to women whose husbands smoke more than 20 cigarettes a day are of lower birth weight than children born to women whose husbands do not smoke.

4.6 Study sample

4.6.1 Target population and accessible population

An important issue in the design of the research is the question of sampling. Ideally, the study design should include all the target population. The term population in scientific methodology refers to the material of the study, whether it is human subjects, animals or inanimate objects. Including all the target population is generally not possible, because of the large numbers, the cost and the time. A subset of the population is studied instead, from which conclusions (or inferences) are drawn as applying to the target population. The sample has to be selected to be as representative as possible of the target population, and in enough numbers to provide valid answers.
The population census is an example of a study in which all members of the population are studied. Even in a small country, it is a very major undertaking. Because of its expense, it is normally carried out every 10 years or so. It normally takes several years to analyse the results. Some countries do an interval census based on subsets of the population in between.

An illustrative example of sampling from another field is that of polls before parliamentary or presidential elections where specialized agencies make predictions based on a relatively small sample representative of the population. Since opinions of voters vary with time before the election, these samplings are commonly done periodically. On the day of the election, samples of exit polls are often accurate in predicting the outcome of the election.

Instead of the “target population”, the investigator often depends on the “accessible population”. The accessible population must be representative of the target population, in order to draw conclusions about the target population. If we take the above example of voter opinions, a polling agency may use the telephone book as the accessible population from which the sample is drawn. This will be acceptable in a country where practically all people have telephones. It will not, however, be representative in a country where a large segment of the potential voters are not reachable by telephone. This does not necessarily mean that the polling should not have been done in this way. The result, however, should be presented as reflecting the opinion of a segment of the target population who are accessible by phone, and not necessarily representing the whole target population.

In health research, the clinic or hospital may provide the accessible population. This, however, does not necessarily represent the community if not everyone goes to the clinic or hospital for the condition in question. This does not mean that clinic or hospital studies should not be done. They provide useful information but the results should not be presented as reflecting the results for all people who have the condition.

### 4.6.2 Types of sampling

The sample selected from the accessible population should be representative of the accessible population. It should accurately reflect the characteristics of the population from which it is drawn. It should be a miniaturized representation of the accessible population.

Random sampling is not haphazard sampling. It is sampling done in a systematic way to ensure, as far as possible, complete objectivity in the selection of the sample. Random sampling is a way of ensuring that all members of the population have an equal chance of being selected. It does not guarantee that the sample will not be different in characteristics from the accessible population. Rather, it eliminates a possible reason that they should be different.
As discussed in section 4.3, random assignment is important when two interventions or more are compared. It minimizes group differences due to biased selection. Randomization was commonly done manually using a table of random numbers. Now, it is usually done using a computer program.

Stratified random sampling is a special type of sampling to ensure that all subgroups in the accessible population are represented in the sample. This is particularly important if certain subgroups are present in small numbers in the population, or are important to be included. In stratified random sampling, key subgroups are defined, for example by sex, social class, income groups, geographic locations, etc. and samples are drawn at random from each of these “strata”. The computer program can be adjusted to draw disproportionately from one or more groups, to ensure their adequate representation.

Cluster sampling is another way of random sampling. It is based first on the random selection of certain subgroups, from which the sample can be taken. For example, in a community survey certain streets or blocks are selected at random first. Then a random sample is selected from each randomly selected cluster. In a health services study, a number of districts are randomly selected. Then a random sample of health service units is selected from each.

Systematic sampling is done by a simple periodic process, for example selecting every second or third patient.

Consecutive sampling involves taking every subject who presents herself/himself over a specified time period. These are not strictly random techniques, but they avoid bias in the selection.

4.7 Sample size

The desired sample size is now easily calculated with the help of computer statistical programs, but the principles underlying the calculation, and the limitations must be clearly understood by investigators.

It is not necessarily true that the bigger the sample, the better the study. Beyond a certain point, an increase in sample size will not improve the study. In fact, it may do the opposite, if the quality of the measurement or data collection is adversely affected by the large size of the study. It is also better to ensure that the sample is representative, rather than being very large.

The statistical concept behind calculation of the desired sample size is simple. When we study a representative sample, we aim to generalize from the sample findings to the population from which the sample was drawn. We cannot be completely certain about this. Unless we study the whole population, the sampling error cannot be brought down to zero. Analytical statistics helps us to define the degree of probability that a finding, a
difference or a relationship can be generalized to the population from which the sample is drawn. This is called the statistical significance of the finding. The size of the sample is an essential element in making this statistical probability calculation. The smaller the size of the sample, the less likely that the findings can be generalized. For calculating the desired sample size before beginning the study, we do the exercise in reverse. We decide beforehand on a level of probability or uncertainty that we are willing to accept for the study, and then we find the desired sample size to provide that level of statistical probability. Traditionally, most studies set this level of statistical significance at 0.05, that is accepting a chance of 5% of finding an association that is not actually there. It must be recognized, however, that this value is arbitrary, and other values can and are sometimes used. In general, the investigator should aim for a lower probability of error when it is particularly important to avoid making a false-positive statement about a finding.

When the study is designed to find a difference or an association, we may not find a difference or an association. In this case, we still want to calculate statistical probability that we may have missed a difference or an association that exists in the population, but was not found in the sample. This so-called statistical power of the study depends also on the size of the sample. The larger the size of the sample, the higher the power of the study. For calculating the sample size before the study begins, the investigators have to make a decision on the level of statistical power they are willing to accept for the study. Traditionally, most studies set statistical power at 0.80, which is accepting a 20% chance of missing a difference or an association that is actually there. It must be recognized, however, that this value is arbitrary, and other values can and are sometimes used. In general, the investigator should aim at a higher statistical power when it is particularly important to avoid false-negative error.

Although a statistician may do the necessary exercise to determine the sample size, s/he can only do it with guidance from the investigator on the level of uncertainty that is considered acceptable. In addition, calculation of the statistical significance and statistical power has to take into consideration some characteristics of the data. These characteristics will thus also be needed for calculating the sample size. Since the data are not available before the study begins, the investigators will have to make some assumptions about the data, and provide these assumptions to the statistician to be able to calculate the desired sample size. The procedure for estimating sample size is not as precise as investigators may be led to think. One such assumption is about the prevalence, incidence or frequency of the condition or event. If the rate of the event is large, statistical power will be high with a smaller number of cases. If the event is rare, a larger sample size will be needed. Also, the larger the variation in the data, the larger the sample size that will be needed to achieve a certain level of statistical significance. For sample size to be calculated, we thus need to make a prior estimate of the frequency of the condition under study, and the degree of variations in the data. Some information may be available
from previous studies to guide the estimates. If not, it is up to the investigators to come up with a tentative estimate which the statistician can use.

The effect size in a study refers to the actual size of the differences observed between groups or the strength of relationships between variables. The likelihood that a study will be able to detect an association between a predictor and an outcome variable depends on the magnitude of the association we decide to look for. Large sample sizes are needed to detect small differences. The choice of effect size is difficult and arbitrary, but it must be set beforehand and must make a meaningful difference. The rule is that the smaller the difference you wish to detect, the larger the sample size needs to be. In designing a study, the investigator chooses the size of effect that is considered important.

In making the final estimation of the sample size, factors such as dropouts, attrition and loss to follow-up should also be accounted for. If the calculated sample size proves to be larger than can be practically obtained, the investigators have a number of options: to increase the effect size they look for; to decrease the power of the study; to modify the design; or to give up the study.

4.8 Measurement

An important question in the research design is the decision on how measurements are made to ensure reliability and validity. Reliability means that the observer repeating the test, or someone else using the same method should be able to obtain the same findings. Validity means that the measurement should actually represent what it is intended to measure.

To ensure reliability or reproducibility of the results the following should be considered.

- Measurements made should not vary by observer or between observers (intra- and inter-observer consistency).
- Instrument or laboratory variability should be taken into consideration.
- Subject variability should be considered if measurements vary according to the time they are made, for example, fasting or after meal, time of the day, or day of the menstrual cycle.

Intra-observer and inter-observer or rater reliability are important issues in measurement. In a study to document them, 29 biopsy slides with suspected Hodgkins disease were presented to three pathologists over an 11-month period (Coppleson et al., 1970). The specimens were unlabelled and over the year of the study were presented on two occasions to each of the three observers. The three observers disagreed with themselves on seven, eight and nine occasions, out of the 29. Overall inter-rater
agreement was calculated at 76% or 54%, according to the particular diagnostic feature described.

Obtaining the same result by the same and different raters ensures reliability and reproducibility, but does not mean validity. The test, itself, may not be accurate in measuring what it is intended to measure. This is particularly apparent in diagnostic tests, as will be discussed in more detail in Chapter 9. The test may be sensitive in detecting people with the disease, but not very specific in excluding people without the condition, or vice versa. To test for validity of the measurement, it has to be compared to a “gold standard”. If for example, we are using a diagnostic test as an indicator of breast cancer, it should be compared to the gold standard of a breast biopsy.

4.9 Planning qualitative research

The above sections dealt with planning quantitative research. Qualitative research needs other approaches (Ulin et al., 2002).

One way to keep the design focused on the research problem is to develop a conceptual framework. A conceptual framework is a set of related ideas behind the research design. A conceptual framework helps to outline the research questions, and provides a context for understanding the research.

Three main methods are commonly used in qualitative research: observation, in-depth interviews and group discussion. The investigator has to select which method would be more appropriate to answer the research question, or may use more than one method. The researcher in these different designs plays the role of observer, interviewer or group moderator.

Observation

Depending on the objective of the study, observation can be made from an outsider or insider perspective, or somewhere in between. Outsider observers maintain a distance. Insider observers interact.

As an example of an outsider observation study, the investigator may observe the quality of health care delivery in a clinic, health centre or a pharmacy. A special type of observation study, called “time and motion study” is used to study how health workers use their time. The researcher observes what a health worker is doing over a defined sample of time. S/he may use a beeper that goes off every number of minutes and a checklist to record activities.

A special form of observation is the so-called “mystery client” technique. It is used particularly in client–provider studies where the presence of an outside observer might change the provider’s customary behaviour. Trained data collectors present as simulated
clients. The deceptive nature of this technique raises ethical concerns. The decision to use the technique should be made only after careful reflection on the ethical implications. Informed consent may be obtained from the health service to use the technique at unannounced times over a period of time, for example several months.

In participant observation, the investigator interacts. S/he may, for example, ask clients about their perceptions of the health service.

**In-depth interviews**

Intensive one-on-one interviewing is a classical method in qualitative research. Different from quantitative studies based on a structured questionnaire, the in-depth interview is more of a social encounter, with questions flowing from the answer of the respondent, as a follow-up to the answer, or to probe further into the answer. Open-ended questioning is a basic tool in qualitative research. The interview may take the form of an informal conversation with little or no preparation and sequencing of questions. Alternatively, a topic guide or outline may be used to help in focusing the interview, but without pre-structuring the questions. A pre-determined set of open-ended questions is, however, the most standardized approach for in-depth interviews.

**Focus groups**

Focus group discussions are the method used when information and insights will be better gained from the interaction of a group than from in-depth interviews with individuals. The two methods may complement each other. A focus group discussion is not a group interview. It is based on the exchange of information, ideas and views among the participants themselves. The researcher is playing the role of a moderator, and not an interviewer. In recent years, focus group methodology has been increasingly used. Certain guidelines need to be observed.

The group should be relatively homogeneous, for example in age and sex and sociocultural background. Anonymity among participants may be desirable, if people feel more comfortable to talk freely with strangers than with people they know and will meet again.

For most purposes, groups of eight to ten participants are adequate for a good and manageable discussion. As to the number of groups, it is generally advised to have at least two groups for each defining demographic variable. If, for example, sex is the variable, two women and two men groups will be needed.

A two-hour discussion is likely to generate 25 to 40 pages of transcript. The role of the moderator is to create a comfortable climate for open exchange, stimulate discussion, keep the discussion focused, and encourage everyone to participate. The moderator should not allow one or two vocal individuals to dominate the discussion.
The rapporteur or note-taker should be recording what people say, but should also be aware of body language.

### 4.10 A note on questionnaire design

A questionnaire is a document designed for the purpose of seeking specific information from the respondents.

The questionnaire may be self-administered or administered by interviewers. The self-administered questionnaire approach is cheap, less susceptible to interviewer bias and can be administered by mail. At the same time, the rate of non-response may be high, and may bias the results. Also, answers may be incomplete.

There are two major question formats: the open-ended and closed-response types. In a closed-response question, the respondent is provided with a list of pre-determined response options. Open-ended questions elicit more detailed responses, but the responses require more effort to encode for data analysis. A questionnaire may include both question formats.

Closed-response questions may be used to elicit attitudes of the respondents to a certain statement. Two formats can be chosen (Polgar and Thomas, 2000). In the Likert-type format, the respondent chooses from among: strongly agree, agree, undecided, disagree, strongly disagree. In the forced-choice format, responses are limited to: strongly agree, agree, disagree, and strongly disagree. This format does not allow an undecided answer.

Questions should be well worded to avoid any ambiguity. Jargon should not be used. Questions should not be phrased in a way that influences the response in one direction or another. The questionnaire should always be pre-tested in a pilot study before the main survey. Interviewers should be trained to make sure that the questionnaire is administered in a uniform way.

A questionnaire typically includes the following components:

- an introductory statement by the interviewer to introduce herself/himself and explain the purpose of the questionnaire; the respondents should also be informed about the confidentiality of their responses;

- demographic questions to collect relevant information about the background of the respondent;

- factual questions;
• opinion questions: opinion questions require reflection; it is generally easier for the respondent to answer factual questions; putting the factual questions first serves as a “warm up” to the opinion questions;

• closing statement by the interviewer to thank the respondents, and where appropriate to ask if s/he wants to provide any additional comment.

A method commonly used to test for reliability in results obtained by questionnaires is to look for internal consistency, that is the extent to which the responses on different questions correlate with each other. If they tend to be highly correlated with each other, then the test is said to be internally consistent. The computer programme can be built up to detect inconsistency.

There is a tendency among investigators to put too many questions. This has been encouraged by the introduction of computer-assisted analysis. Information collected in a questionnaire should be based on and limited to the objectives of the study.

4.11 A note on research in health economics

All methods of economic evaluation in health care have one principle in common: they examine one (or more) possible interventions and compare the costs of inputs or resources necessary to carry out such interventions with their effects or economically assessed benefits (Jefferson et al., 2000).

In economic evaluation, the cost of an illness generally includes:

• direct costs, which are costs borne by the health care system, community and patients’ families in addressing the illness (for example, diagnosis or treatment costs);

• indirect costs, which may be tangible or intangible; indirect tangible costs are mainly productivity losses, caused by the disease condition, and borne by the individual, family, society, or by the employer; indirect intangible costs include the costs of pain, grief and suffering, and the loss of leisure time.

In economic evaluation, resources are estimated as all inputs into health service production, including time, goods, equipment, buildings, specialized knowledge, etc.

Cost-benefit analysis and cost-effectiveness analysis are related analytical methods that compare health care practices or techniques in terms of their relative economic efficiencies in providing health benefits. In a cost-effectiveness analysis, the net monetary costs of a health care intervention are compared with some measure of clinical outcome or effectiveness, such as cases of disease avoided, cases identified in screening procedures, life years gained, or deaths avoided. Cost-benefit analysis compares monetary costs to estimated monetary benefits of an intervention.
Cost-effectiveness analysis is frequently nested within a randomized controlled trial. It is particularly valuable when the compared interventions have widely differing costs or resource consequences. Competing interventions in the trial may show little difference in outcome. The addition of the economic perspective offers a further dimension of evaluation. Prospective economic data collection alongside a trial allows the evaluation to be based on reliable estimates of effectiveness.

4.12 Ethics in research design

4.12.1 Categories of health research

From an ethical standpoint, four categories of health research can be distinguished.

- Research involving human experimentation: This is the research category that raises most ethical concerns. Under this category, two types of medical research can be distinguished: a) research of therapeutic or diagnostic nature that is carried out on patients who may expect a potential benefit from their participation; and b) research of a purely scientific nature for which human subjects volunteer to advance medical science but will not draw any therapeutic or diagnostic benefit. Ethical safeguards are most needed in this category.

- Research involving human subjects but not experimentation: Epidemiological and field studies, as well as qualitative research, fall under this category. Although no experimentation is involved, such studies can be intrusive on the individual’s privacy and even on communities.

- Research involving experimentation on animals: Ethics in this category has been receiving increasing attention recently.

- Research not involving human subjects or animal experimentation: This category of research would still be bound by ethical principles that cover research in general, medical and non-medical.

4.12.2 Ethics in research design involving experimentation on human subjects

All research involving human subjects should be conducted in accordance with the ethical principles contained in the current version of the World Medical Association Declaration of Helsinki (Annex 1). All individuals involved in the conduct of any clinical trial must be fully informed of and comply with ethical principles, including beneficence, non-maleficence and respect.
The principle of beneficence implies that:

- a scientific and technically sound design is an ethical requirement; a design that will not provide the answer to the research question is ethically unacceptable, as the patients will be subjected to an unnecessary process;
- the sample size is adequate to provide statistically valid results, but is not larger than is necessary to provide the answers.

The principle of non-maleficence implies that:

- any potential risks are properly evaluated and balanced with potential benefits, are minimized in every way possible, including adequate screening for contraindications, and are carefully monitored;
- where adverse effects are encountered, adequate treatment is provided.

The principle of respect implies that:

- participants are fully informed and give their free consent to participate in the trial;
- research trials on children and persons with mental disability are limited to disease conditions specific to them and the informed consent of parents or a guardian is obtained;
- confidentiality is adhered to.

Confidentiality is an ethical obligation in the practice of medicine. Since in research, information is likely to be handled by other people involved in the research, steps should be taken to ensure the confidentiality of the records either by limiting access or by replacing patient identification with code numbers.

A number of ethical considerations apply when a new therapy is being tested on patients, according to the principle of “do no harm” or non-maleficence.

- Pre-clinical studies that provide sufficient documentation of the potential safety of the pharmaceutical product should be available.
- Information about manufacturing procedures should establish that the product is of suitable quality.
- The data available should be appropriate to the phase, size and duration of the trial.
- Data from previous and ongoing clinical trials should be compiled before the trial.
- The investigators should be well qualified and the trial site adequate.
• All parties involved in a clinical trial should comply fully with the existing national regulations or requirements.

4.12.3 Epidemiological, field and qualitative studies

This research is based mostly on observation, and generally requires no intervention more invasive than asking questions and carrying out routine medical examinations and, sometimes, laboratory tests or X-ray examinations. Such studies do not carry physical risks for the research subjects. However, they can be intrusive. Psycho-social harm may be as or more meaningful to the person than physical harm. Ethical considerations include free informed consent, confidentiality and beneficence.

The principle of free informed consent implies that individual subjects should understand and agree to the reasons for collecting the information. In large community surveys, the community must also agree to the study.

The principle of confidentiality implies that information gathering in qualitative research is based on mutual trust. This trust will be seriously breached by any possibility of break of confidentiality. Information collected about subjects in field studies is generally classified as linked or unlinked (CIOMS, 1991). Unlinked information is information which cannot be linked, associated or connected with the person to whom it refers. Confidentiality here is not at stake. Linked information may still be anonymous, if it is linked to the person by a code or other means, and the investigator cannot know the identity of the person. In other cases, strict adherence to confidentiality should be maintained.

The principle of beneficence implies that:

• The individual has a right to be informed of any health condition revealed during the study, and should be helped to get the appropriate care.

• The community has a right to be informed about the outcome of the study, and any potential implications.

• The investigators have the ethical obligation to play an advocacy role to improve the health condition of the community based on the results of the study.

• Local personnel should be utilized, as far as possible, and they should be trained in the required skills. An ethically conducted epidemiological or field study should leave something behind in the community in which it was conducted. So-called “safari research” should be discouraged.
4.12.4 Ethics in research designs involving experimentation on animals

The animal model chosen must be relevant to the human. The information must be applicable to the human.

The minimum number of animals should be used. Experiments should be designed with proper calculation of the size of the animal sample needed to answer the research question or test the research hypothesis. No more than the minimal number of animals should be used, but a sufficient number of animals should be used to provide a scientifically valid conclusion.

References and additional sources of information


Byrne DW. Publishing your medical research paper. Baltimore, Lippincott Williams & Wilkins, 1998: 5–44.


