

Clinical Research Methods

Primer of Epidemiology V: Planning a research study and sampling methods

ROOPA SHIVASHANKAR, KAVITA SINGH, PRITI GUPTA

DEVELOPING A RESEARCH QUESTION

The first critical step in the planning of a study is to develop an answerable research question. A well-crafted research question is key to designing and conducting epidemiological research. A poorly developed, less thoughtful research question may result in a haphazard research study, which may not prove useful in addressing a knowledge gap. In this article, we discuss the processes involved in writing a research question and features of a good research question.

Often, one may become curious about a potential issue owing to observations in the clinic, exchanges during a lunch discussion with colleagues or while reading the literature. Initially, the query may be vaguely defined. One such example is given below.

A physician observes that a large number of patients in weekly hypertension clinics have uncontrolled hypertension during the winter months. She/he wishes to identify strategies to improve blood pressure (BP) control. This idea needs to be refined to define a proper research question by dissecting it into smaller research ideas. In this example, many different dimensions can be studied:

- Is there a true surge in uncontrolled hypertension in the winter months?
- Which groups are more likely to have uncontrolled hypertension? Elderly? Men? Patients with comorbid conditions?
- Is the rate of uncontrolled hypertension in the clinic similar or different from that of other clinics?
- Are there factors in the environment, such as festivals, that may be driving this apparent surge?
- What advice can be given to patients to improve control during winter months?

These directions will help the researcher think through the idea and look for answers in the existing body of literature. The researcher may find that the clinic has similar rates of uncontrolled hypertension as other clinics. Uncontrolled hypertension is likely to be due to lack of treatment adherence and that clinics

in developed countries have used counselling techniques to improve adherence to treatment. With this process, the researcher may arrive at a crude research question as follows: Would having patient counselling for adherence to treatment at the hypertension clinic improve the treatment outcomes at weekly clinics?

This question needs to be further refined by using specific definitions. A research question should consist of information on participants, intervention, comparison, outcomes and study design (PICOS).

In the above example, using the PICOS approach, a more specific research question can be framed:

- *Population*: Should the investigator recruit any patient with hypertension or only uncontrolled patients with hypertension? Should children be included?
- *Intervention*: Counselling for adherence to treatment needs more clarity. Who will do the counselling—physician or nurse or a qualified counsellor? How long and how frequently will the counselling be carried out?
- *Comparison*: Is there a comparison group? If so, what is the comparator group/control arm?
- *Outcomes*: What factors (primary or secondary) will be assessed or measured to see the effect/change in clinical parameters with the implementation of a new treatment strategy? Is it BP control? Reduction in complications? When is the improved treatment outcome measured? At the end of 3 months, 1 year or 2 years?
- *Study design*: Will this be an observational study of the existing programme or an intervention trial? If it is the latter, is randomization done at the patient or clinic levels?

By thinking through these questions, the research question could be refined to make it more specific as given below:

Do adult patients with systolic BP >180 mmHg at the hypertension clinic randomized to receive an additional 10 minutes of treatment adherence counselling have improved BP control (<140 mmHg) at the end of 6 months compared with those who received usual counselling by physicians?

A good research question will be clear and specific and includes information on PICOS. Researchers should spend sufficient time thinking through a research idea. It is worth discussing the research question with colleagues/friends and receiving feedback on its clarity, novelty, practicality, relevance and ethical concerns.

DECIDING ON AN APPROPRIATE STUDY DESIGN

In the previous articles, we discussed various designs used in epidemiology and their strengths, drawbacks and applications. We will discuss how to choose an appropriate design to

Centre for Chronic Disease Control, C-1/52, Safdarjung Development Area, New Delhi 110016, India
ROOPA SHIVASHANKAR, PRITI GUPTA

Centre for Chronic Conditions and Injuries, Public Health Foundation of India, Gurugram, Haryana, India
KAVITA SINGH

Correspondence to ROOPA SHIVASHANKAR; roopa@ccdcindia.org
[To cite: Shivashankar R, Singh K, Gupta P. Primer of Epidemiology V: Planning a research study and sampling methods. *Natl Med J India* 2021;34:287–92.]

© The National Medical Journal of India 2021

conduct a particular study or answer the proposed research question. Broadly, the choice of the study design depends on the research question. However, many other practical aspects such as availability of resources (human, physical and funds) and time; the feasibility of recruitment of eligible patients/study participants and follow-up, practicality, logistics and ethical issues also play a role in deciding on the research question.

Table I provides some broad guidelines to choose study designs for different types of research questions. The choice of design for descriptive studies is straightforward. However, some aspects are needed to be considered in choosing designs for explanatory studies. If the research is dealing with rare diseases (e.g. congenital heart disease), limited funds and time availability, then case-control studies are the design of choice. However, if the research question aims to answer a temporal association of exposure with a common disease (e.g. salt intake and incidence of hypertension) with sufficient funds and time, then a cohort study would be the right option.

Conventionally, experimental designs, specifically randomized controlled trials (RCTs), are ranked higher in the hierarchy of strength of evidence (Fig. 1). Experimental designs are particularly useful to assess the effects of interventions (new treatment). However, it is not always feasible, practical, ethically correct or scientifically appropriate to do an RCT. For instance, it is ethically incorrect to randomly assign a possible risk factor (e.g. chewing tobacco) to study participants. However, experimental designs can be used to assess the effect of removal/reducing the uptake for risk factors (e.g. tobacco cessation and salt reduction).

DATA COLLECTION METHODS/TOOLS

There are many methods of collecting data for research purposes. A single variable can be obtained by many methods. For instance, if 'hypertension' is a variable of interest in a particular study, the data on hypertension can be obtained from one or more combinations of 'self-report of participants', 'reviewing medical records' or 'objective measurement of BP'. The BP can be measured with either a manual sphygmomanometer or an electronic BP-measuring device. Further BP readings can be taken once, multiple times or continuous 24-hour monitoring. The choice of the method depends on the purpose of the study, validity and reliability of the method, availability of resources and feasibility and acceptability to the study population.

We now describe the common data collection methods/tools, their uses and their limitations. The data collection

methods can largely be stratified into questionnaire methods and objective measurements.

QUESTIONNAIRES

Questionnaires are sets of questions with fixed response categories designed to obtain data on personal (demographic and socioeconomic) characteristics, exposure factors, confounders and sometimes outcome variables. They are a common method of data collection as they are relatively less expensive, and objective methods are not always available or feasible or affordable. Data about past exposures can be obtained only by a questionnaire unless good biomarkers are available.

Questionnaires can be either administered by in-person interview or self-administered:

- In-person interviews have historically been the most common method of data collection in epidemiology. They are the most advantageous. This method has higher response and completion rates. It allows the inclusion of illiterate participants and the use of a lengthier and complex questionnaire and permits clarification and explanation. Nevertheless, this method does have some disadvantages. It can have social desirability bias (participants less likely report socially unacceptable behaviours if asked by an interviewer compared with self-report). Also, the cost of conducting interviews is higher compared with that of self-administered questionnaires.
- Self-administered questionnaires are relatively cost-effective and have less social desirability bias. However, people are less likely to respond to a self-administered questionnaire specifically when sent by email or post. Moreover, the low completion rate is also an issue due to the absence of the

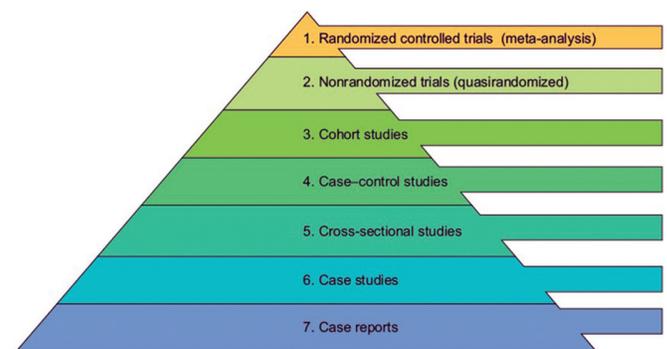


FIG 1. Hierarchy of study designs

TABLE I. Guidelines to choose study design to suit the research question

| Types of research question/study | What is measured? | Example | Choice of study design |
|----------------------------------|-----------------------|--|-----------------------------|
| Descriptive | Prevalence/burden | Prevalence of obesity in secondary schoolchildren; prevalence of uncontrolled hypertension in primary care centres | Cross-sectional |
| | Incidence | Incidence of secondary myocardial infarction in the hospital | Cohort* |
| Explanatory-observational† | Association/causation | Does high salt intake cause hypertension? | Case-control or cohort |
| Explanatory-experimental | | Does lowering salt intake reduce blood pressure? | Randomized controlled trial |
| Exploratory | Attitude/behaviour | Why do some women choose to do leisure-time physical activity and others do not? | Qualitative methods |
| | Reasons/explanations | Why did the quality improvement in the cardiac care unit fail? | |

*Registries are special clinical cohorts that are useful in assessing the incidence of complications, side-effects of drugs and mortality can be used for analytical studies, they are hypothesis generating rather than hypothesis testing

†Although cross-sectional design

interviewer to prompt or motivate. Self-administered questionnaires are not an option if a large proportion of the target population is illiterate. It is not a feasible method for complex questionnaires. Of late, with higher accessibility to electronic devices such as mobile phones, tablets or computers, self-administered questionnaires can also be provided through the internet. There are many applications and websites that host questionnaires. These are useful as they can reach larger samples at lower costs. Large-scale access to computers and the internet is, however, the prerequisite for this method.

- *Developing a questionnaire:* Questionnaires should be designed and tested well before the study. Generally, questionnaires are designed in English and, in the Indian context, need to be translated into local languages, which need additional time. It is a good idea to search for standard questionnaires that are pre-validated in the Indian context, which may meet the whole or part of the objectives of the study. This would not only reduce the time and efforts of questionnaire development, but also make it easier to compare the results with other studies.

While designing a questionnaire, one has to ensure that it meets the objectives of the study, is easy to comprehend for both the interviewer and the interviewee, has a lower respondent burden, is culturally sensitive, has minimal processing requirements and has low measurement error.

Questionnaires should contain sufficient information that meets the objective of the study. One needs to be careful about adding questions not directly relevant to the study. Although it is opportunistic to collect more information at one visit that may be useful for other research questions, this would increase the respondent burden, which may lead to lower completion rates and low quality of data. The recommended maximum interview time with participants is between 40 and 60 minutes.

OBJECTIVE ASSESSMENT METHODS

Objective assessment methods do not require the participant's report and are usually based on examination or clinical or laboratory measures. They have advantages over the questionnaire methods, as they overcome personal bias and recall errors. For example, participants may either deny smoking history or not accurately recall the amount of tobacco consumption. Measuring cotinine content in saliva or urine can overcome such inaccuracies and more validly capture the exposure to tobacco.

Objective assessment methods are constrained by factors such as availability in the local setting, accessibility of validated instruments, their feasibility in the field and affordability. For measurements that rely on blood samples, for example, there may also be lower response rates due to reluctance to agree to the procedure, thus introducing potential selection bias at the time of analysis. In addition, non-invasive procedures such as obtaining saliva increase the time burden of both the respondents and investigators in the field. Investigators have to give clear instructions on how participants should provide their samples and must arrange for another visit to get the samples. Participants have to follow the correct steps and remember to provide the sample in the morning. This may also lead to a lower response rate. Furthermore, cotinine assays are expensive. Therefore, a great deal of judgement is required before deciding on the method of data collection.

DEVELOPING A STUDY PROTOCOL

The study protocol is a document that describes the background, aims and objectives and methods to be followed in the proposed study. This document helps the researcher in putting their thoughts together; it also conveys clear instructions for reproducible procedures, a cornerstone of a sound scientific method. A study protocol ensures transparency with funding agencies and collaborators. Many large research studies publish their protocol in peer-reviewed journals. The common structure of a study protocol is as follows:

Title

The title should convey the gist of the study in a phrase or sentence. Usually, it will be between 8 and 15 words. A good title will consist of PICOS and also place or site of the study. Some examples are:

- Effect of a triple pill on hypertension compared with usual treatment among the elderly in community health centres of Haryana: A double-blind randomized trial
- A cross-sectional study of the prevalence and correlates of tobacco use in Chennai, Delhi and Karachi: Data from the CARRS study
- Adherence to diabetes care: A cross-sectional survey of processes at general practices in NCR Delhi, India

Background

This section should briefly explain the importance of the health issue the researcher plans to address—from what is already known about the health issue both globally and locally, to what the knowledge gaps are and how the current proposed study plans to address this knowledge gap. It may be a good idea to describe how the chosen population, study design and setting help in the knowledge enhancement about the concerned health issue. This section should build an argument for doing the current research study and outline the potential usefulness of its results.

Aims and objectives

'Aims' is the broader description of the overall goal of the study and is a general statement. For example, a study may be aiming to measure the burden in chronic kidney disease (CKD) in urban India. The aim may be written as follows: 'To measure the burden of CKD in urban India'.

However, an objective is a more specific description of the research question. It should have a measurable outcome. PICOS should be included in the objectives. The objectives of the same study that aimed at measuring the burden of CKD in urban India could be written as: 'To estimate overall and age-, sex-, city- and diabetes-specific prevalence of CKD using a standardized definition from the 2012 Kidney Disease International Global Outcomes CKD among adults in two major cities in India (Delhi and Chennai)'.

Methods

This section describes the study design, study population, setting, sampling, sample size, measurements of exposure, outcomes and other confounders, ethical issues, quality control, etc. In other words, this section will provide details on how the research question will be answered. The subsections will depend on the nature of the study design. The common sections for all study designs are described below.

The study design should include a description as to whether the study will be prospective or retrospective for a cohort study and whether there will be single/double/triple/no blinding for an RCT.

Study setting. This section describes the setting of the study in terms of its geographical location, whether it will be rural or urban, in the community or a hospital.

Study population. This section includes who would be the participants in the research (inclusion and exclusion criteria) and how they will be chosen (sampling methods). This section describes the participants who would be studied in the research in terms of age group, sex, any specific disease or risk condition, etc. For a case-control study, the characteristics of cases (new diagnosis or prevalent cases, with complications or not, etc.) and controls (hospital- or community-based) and matching criteria, whether pair or group matched, etc. are described. For each group that will be included, the screening-out process in the exclusion criteria should be mentioned. For example, the researcher may want to exclude extremely sick or pregnant women.

Sampling methods (details in the next section). What sampling method will be used? Why was this method chosen? What steps will be taken in case of refusal and non-consent or non-participation?

Sample size. This section will describe the planned number of participants to be enrolled in the study and the basis for the sample size calculation including assumptions and their references. (Several online tools and calculators are available for sample size estimations. The required sample size can be estimated in consultation with a biostatistician.) If the study involves two or more groups, ensure the planned number of enrolment in each group. It may be a good idea to provide a table for sample size calculation using different assumptions and describe in the text the chosen estimated sample size and why it was chosen.

As discussed earlier, the sample size calculation should account for the sampling design effect and the expected non-response rate. Tabulation of various sample size calculations and references for their basis will help to explain the chosen sample size; Table II gives one example.

Randomization. For conducting an RCT, write the plan for randomization—the type of randomization, block size, level of blinding and execution.

Variables. This section defines and describes how the main exposure(s), outcome(s) and confounders will be measured. Explain why a particular definition and method of measurement are chosen. If a questionnaire will be used, mention whether the questionnaire is validated for the community and the language in which the research will be conducted. In addition, mention whether the questionnaire is self- or interviewer-administered, and who will administer the questionnaire (pre-required education, experience and training offered)

For the other methods, mention which instruments will be

used and whether they are standardized. Who will do the measurements? How will intrapersonal and interpersonal variability be assessed? What are the plans for the calibration of instruments used in the study?

In follow-up studies that involve more than one contact with the participants, describe what measurements are done in the baseline visit and subsequent follow-up visits.

If the study involves more than one site, include how the uniformity of data collection is ensured and provide details on quality assurance and quality control plans.

Ethical issues. How does the research ensure that the rights of research participants are protected? What steps are taken to ensure the confidentiality of research data? What are the plans to protect vulnerable groups (children/pregnant mothers)? What is the process of informed consent from participants? Is the protocol approved/submitted to any ethics committee?

Data management and analysis plan. It is strongly advisable to involve a biostatistician(s) during protocol development. With the help of biostatisticians, develop a data management and analysis plan. Plan how the database will be developed for data entry (platform, outlier and logic checks) and methods to reduce data entry errors (e.g. double data entry).

Prepare an analysis plan during protocol development for the primary objectives of the study. With the help of the statistician, make a plan of how the data will be summarized and statistical and regression methods that will be used for comparison of groups. Make a note of this plan in the study protocol.

References. Include all relevant references for background, methods, sample size estimation and analysis plan in the standard format.

Appendices. This section will have supplementary materials that are too large to be included in the protocol, not essential but useful to understand the study protocol. This may include a questionnaire (or a draft of a questionnaire), a table of various sample size estimation, participant information sheet and informed consent form.

As further guidance, reporting guidelines that have been developed for preparation of manuscripts may be useful in the development of a study protocol. For example, the Consolidated Standards of Reporting Trials (CONSORT) guidelines provide parameters for reporting on RCTs, whereas the STrengthening the Reporting of OBServational studies in Epidemiology (STROBE) guidelines provide parameters for reporting on observational studies. Such guidelines contain a minimal set of considerations for developing and reporting a study protocol.

SAMPLING METHODS

Why is sampling needed?

Sampling is done on the basis of costs and other logistical issues as it is not possible to study all the patients in a population. Thus, sampling is done to obtain data from a sample population so as to be able to extrapolate the results to the whole population.

TABLE II. Example for a sample size estimation: Different levels of prevalence, error rates and response rates for measuring the prevalence of non-alcoholic fatty liver disease in urban New Delhi

| Reference and assumptions | Level of confidence | Error | Prevalence | Design effect | Response rate | Estimated sample size |
|---|---------------------|-------|------------|---------------|---------------|-----------------------|
| Amarapurkar <i>et al.</i> ¹ Mumbai (urban) | 1.96 | 0.05 | 0.16 | 1.5 | 0.8 | 387 |
| Mohan <i>et al.</i> ² Chennai (urban) | 1.96 | 0.05 | 0.32 | 1.5 | 0.8 | 627 |
| Das <i>et al.</i> ³ West Bengal (rural) | 1.96 | 0.05 | 0.10 | 1.5 | 0.8 | 259 |

Steps in sampling

- Defining the population (before a sample is taken, the population is defined to which the results will be generalized or extrapolated; this is called the sampling frame; it is necessary to have a complete list of the sampling units in the target population; the list must be mutually exclusive, recent and exhaustive)
- Determining the sample size
- Drawing the sample
- Starting the survey

Types of sampling methods

- *Probability sampling:* The sampling unit has a known probability of being selected.
- *Non-probability sampling:* The sample does not have a known probability of being selected.

Probability sampling

Simple random sampling (SRS). This is the simplest and best form of sampling technique. A random selection does not mean a haphazard selection. In epidemiology, random selection means that the sampling units in the population have equal chances of being selected.

The most common form for random selection used in day-to-day life is the draw of lots. In this method, the name or identifier of every eligible participant is put in a box and the required number of participants are drawn from the lots. Other methods include using a table of random numbers and most commonly, computer-generated random numbers.

Advantages

- Easy to understand and can be quickly implemented
- Prior knowledge about the population not needed
- Simpler to use statistical methods in the analysis as most methods assume simple random selection (there is no need to adjust for sampling errors).

Disadvantages

- SRS requires a prior sampling frame to draw the sample from. This is logistically difficult and expensive.
- Selected units can be physically far apart, thereby increasing logistical difficulties.
- In small sample sizes, there may be a skewed distribution of participants due to chance alone.

Systematic random sampling. This type of sampling is much easier to apply in practice and also ensures random representation. In this method, only the first unit is selected randomly; the remaining sample is selected in a systematic fashion after every sampling interval. The sampling interval is calculated as a fraction of a sample over the total population. For example, if one needs to select 200 households (sample size) from a village of approximately 2000 households (total population), then one would need to select every 10th household. Therefore, the sampling interval, in this case, is 10. The first unit will be selected from the first ten households randomly, and every 10th household is recruited subsequently in the study. Let us say that the first house selected was house number 6 and then subsequently 16th, 26th, 36th, ..., houses will be selected.

Systematic sampling is practical and relatively easy to conduct. If done in the right context, this will provide the representative sample. However, this sampling design is prone

to systemic errors. For instance, in the above-mentioned example, if every 10th household happens to be a corner house or ground floor house in multi-storeyed buildings, these householders could be systematically different from the rest of the householders (corner house and ground floor are usually more expensive and richer persons are likely to reside there). Therefore, systematic sampling should be used cautiously.

Complex sampling methods. Simple random sampling is not practical in many situations. In the situation where sampling framework is not available or outdated, it is not feasible. If the population is diverse or spread out, simple methods are either not representative (unless in large samples) or not practical. Complex sampling methods are used in such situations. Stratified cluster sampling and multistage sampling are the types of complex sampling.

Stratified random sampling. SRS cannot ensure that the participants from various strata (age groups, sex and urban-rural) are represented with a sufficient sample size in a study. Some of these factors may influence the outcome of interest. For example, if the outcome of interest is the prevalence of smoking in a particular community, and the sample consists of 20% men and 80% women, this would underestimate the prevalence of smoking. In such cases, SRS ensures equal representation of variables of interest.

In stratified sampling, the population is stratified by a variable of interest (sex/age/area, etc.) into groups (strata). A simple random sample will be taken from each group (stratum).

Cluster sampling. Clusters are a collection of individuals in groups. These could be villages, neighbourhoods, schools, hospitals, wards, census blocks, etc. In cluster sampling, a simple random sample of clusters is chosen and all the individuals in the chosen clusters are studied. For example, if the research aims to study obesity in school-going children in the Bengaluru district, the study will list all the schools in the Bengaluru district and take a random sample of schools and study all children in the selected school.

Cluster sampling is convenient to conduct if the population is spread out. However, the statistical analysis used with cluster sampling is not only different but also more complicated. This is because the individuals within a cluster are more similar compared with other individuals (intracluster correlation) and therefore, sample sizes need to be adjusted for these complex designs (design effect).

Multistage sampling. Multistage sampling (Fig. 2) refers to sampling plans where the sampling is carried out in stages using smaller and smaller sampling units based on a hierarchical structure of the population. This is employed when doing research in large geographical areas. For example, the National Family Health Survey and sample registration system use multistage sampling in India. The first sampling units are called primary sampling units. There are two ways of doing multistage sampling: one, by using SRS of primary sampling units followed by random sampling of secondary units keeping the number of secondary units constant.

The second method is by using population proportion to size sampling of primary sampling units followed by random sampling of secondary units, keeping the fraction of secondary units constant. Probability proportion to size is a sampling procedure in which the probability of a unit being selected is proportional to the size of the ultimate unit, giving larger clusters a greater probability of selection and smaller clusters a lower probability. It is most useful when the sampling units

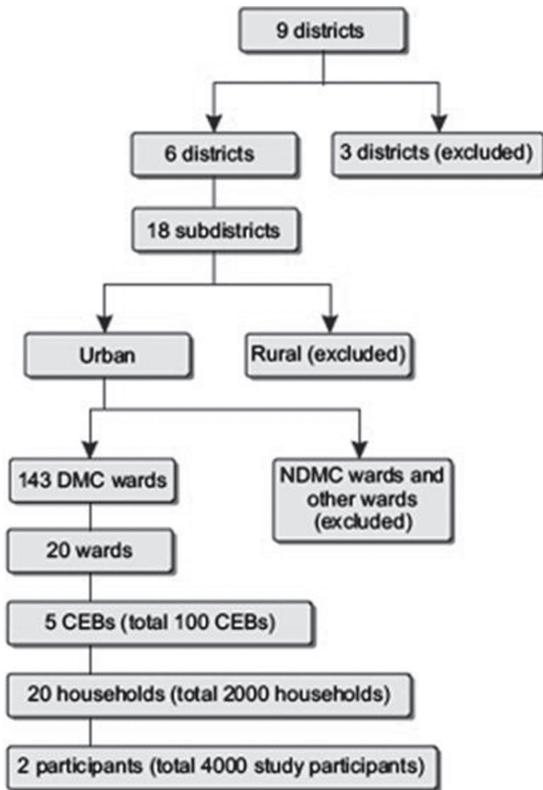


FIG 2. An example of a multistage cluster random sampling (Source: Centre for Cardiometabolic Risk Reduction in South Asia Surveillance Cohort study in New Delhi).⁴ DMC Delhi Municipal Corporation NDMC New Delhi Municipal Corporation CEB census enumeration block

vary considerably in size because it assures that those in larger sites have the same probability of getting into the sample as those in smaller sites and *vice versa*. (For further details, see ‘Suggested Reading’.)

Non-probability sampling

When probabilistic sampling is neither feasible nor required, many non-probability sampling techniques are used. When research needs to be done in unfavourable conditions (minimal budget/dangerous situations), a convenience sampling is done in which individuals more readily accessible to the researcher are more likely to be included. Typically, qualitative research uses purposive sampling, that is, carefully select participants based on study purpose with the expectation that each participant will provide unique and rich information of value to the study. For research that involves a specific population (stigmatized groups such as injection drug users, infective endocarditis, altered sexual orientation and risk for coronary

artery disease), snowball sampling is used. A person with required characteristics is asked to identify other persons with the same characteristics, and those, in turn, are asked the same. These methods are used to understand patient preferences and patient behaviours. For example, Ramakrishnan *et al.*⁵ studied the sex differences in the utilization of surgery for congenital heart disease in India using in-depth interviews (a qualitative research technique) and identified apprehensions about future matrimonial prospects of girls and lack of social support as the major factors responsible for delays in undergoing surgery.⁵ This insight would not have come from a conventional research design.

Thus, these samplings methods are usually not representative and selected to serve the purpose of formative research. Conventional statistical methods cannot be applied in non-probability sampling. A detailed discussion of qualitative research is beyond the scope of this article, and readers are referred to qualitative research methods by Hennink *et al.*

ACKNOWLEDGEMENTS

We are grateful to Dr D. Prabhakaran, Vice President for Research, Public Health Foundation of India (PHFI), for critical inputs in developing this manuscript. We also thank Sanjana Bhaskar, Research Assistant, Centre for Environment Health, PHFI, for editorial assistance.

Conflicts of interest. None declared

REFERENCES

- 1 Amarpurkar D, Kamani P, Patel N, Gupte P, Kumar P, Agal S, *et al.* Prevalence of non-alcoholic fatty liver disease: Population based study. *Ann Hepatol* 2007;**6**:161–3.
- 2 Mohan V, Farooq S, Deepa M, Ravikummar R, Pitchumoni CS. Prevalence of non-alcoholic fatty liver disease in urban South Indians in relation to different grades of glucose intolerance and metabolic syndrome. *Diabetes Res Clin Pract* 2009;**84**:84–91.
- 3 Bajaj S, Nigam P, Luthra A, Pandey RM, Kondal D, Bhatt SP, *et al.* A case-control study on insulin resistance, metabolic co-variables and prediction score in non-alcoholic fatty liver disease. *Indian J Med Res* 2009;**129**:285–92.
- 4 Nair M, Ali MK, Ajay VS, Shivashankar R, Mohan V, Pradeepa R, *et al.* CARRS surveillance study: Design and methods to assess burdens from multiple perspectives. *BMC Public Health* 2012;**12**:701.
- 5 Ramakrishnan S, Khera R, Jain S, Saxena A, Kailash S, Karthikeyan G, *et al.* Gender differences in the utilisation of surgery for congenital heart disease in India. *Heart* 2011;**97**:1920–5.

SUGGESTED READING

- 1 Hennink MM, Hutter I, Bailey J. *Qualitative research methods*. Thousand Oaks, CA: Sage Publications Ltd.; 2011.
- 2 White E, Armstrong BK, Saracci R. *Principles of exposure measurement in epidemiology – Collecting, evaluation, and improving measures of disease risk factors*. 2nd ed. Oxford, UK: Oxford University Press; 2008.
- 3 Smith PG, Morrow RH. *Field trials of health interventions in developing countries: A tool box*. Oxford, UK: Macmillan; 1996.
- 4 Abramson JH, Abramson ZH. *Survey methods in community medicine*. Edinburgh, UK: Churchill Livingstone; 1999.
- 5 Fathalla FM. *A practical guide for health researchers*. Cairo, Egypt: World Health Organization Regional Office for the Eastern Mediterranean Cairo; 2004.
- 6 Bonita R, Beaglehole R, Kjellström T. *Basic epidemiology*. 2nd ed. Geneva: World Health Organization; 2006.