

Epidemiological study methods

Data collection methods

Part I

Health research methods could be divided into qualitative and quantitative methods.

Quantitative methods are those methods which are used to collect information related to quantities such as numbers, frequencies, proportions, effect sizes etc. these methods aim to answer questions such as “how much”, “how many”; “how frequently”, “what proportion”, “how many times”.. We use quantitative methods to collect information on age, education level, serum bilirubin level, percentage of smokers, effect of hyperlipidemia on stroke etc.

Qualitative methods are used to collect information on the meaning of health events and in order to provide understanding and context for these events. Qualitative methods aim to answer questions such as “why”, and “how”. For example these methods can tell us why people fail to wash their hands before eating, how people perceive their health, why diabetics fail to comply with their medical advice and how satisfied users of a PHC are.

There are many methods of data collection within each category of quantitative and qualitative methods. The decision to utilize a particular methods or more than one method is related to the study circumstances. The following questions could help clarify which methods are better for a particular study:

- Which method helps best in achieving study objectives?
- Which methods are appropriate to the design of the study?
- Is the chosen method feasible and practical in the study circumstances?
- Are funds and resources available to use the chosen method?

Generally we use either quantitative or qualitative methods of research but sometimes it may be more useful to combine both methods in a study. Reasons for combing these research methods include:

1. Using qualitative methods in the preliminary stage of a quantitative method to explore the situation (exploratory research) such as when we enter a community we don't know well and need to understand their culture in order to facilitate the research implementation or when we want to find best terms to prepare a questionnaire.
2. Using qualitative methods to supplement quantitative methods to enrich the information obtained. For example in the study of an outbreak of cholera we may need to undertake rapid appraisal of the community or undertake key informant interviews to understand the community.
3. Using qualitative methods to investigate certain behaviors which cannot be investigated properly by quantitative methods. For example, observation may be a better way to investigate hygiene-related behaviors in a community than quantitative interviews.

Main quantitative methods which will be explained in these lectures of include the following:

1. Interviews using a questionnaire
 - a) Face-to-face interviews with research subjects
 - b) Telephone interviews
 - c) Self-administered questionnaires (mail questionnaires)
 - d) Proxy interviews with people related to subjects
2. Existing records
3. Diaries
4. Physical examination and clinical measurement (measurement on subjects)
5. Measurements in the environment

Main qualitative methods explained here include the following:

1. Qualitative interviews
2. Observation
3. Focus groups

Quantitative Data Collection Methods

1) Interviews using a questionnaire

A questionnaire is a written document which is used in research to obtain information from readers or listeners through answering a set of questions (items). A questionnaire is a necessary survey instrument for most research methodologies when information is collected from subjects. Questionnaires may be administered by an interviewer as in personal interviews with the participant, proxy interviews and telephone interviews or they may be self-administered as in mail questionnaires.

The main objectives of a questionnaire are to obtain measurement of exposures and outcomes and to minimize measurement error. These objectives should be kept in mind while designing a questionnaire. In addition a questionnaire should be easy to use, process and analysis and it should also respect the respondent's dignity and privacy.

Standard questionnaires

One must always review the literature before designing a questionnaire and seek help from colleagues and friends who might be aware of previous work in the research area. There might be standard questionnaires or at least standard questions, and if there are, they must be used if appropriate. The advantages of using a standard questionnaire are:

1. They are usually proved to be satisfactory in use
2. Their reliability and validity have been checked before
3. They are good for comparison with other studies
4. They facilitate work.

However, there may be no standard questionnaires for the research we want to do in which case we have to design our own questionnaire.

What does a question measure?

A questionnaire contains a set of questions laid to measure different exposures, confounding variables and outcomes. It is useful to know what kind of information each question seeks to obtain. Questions can be divided to the following types depending on what they measure (the nature of information they aim to obtain):

1. Attribute: what one is. Example
How old are you?
What is your profession? (Open or close-ended answer)
2. Knowledge: what one knows. For example
Which of the following factors can cause diarrhea in children?
 - a) Swimming in a contaminated pool
 - b) Storing drinking water in un-covered containers
 - c) Bottle-feeding
 - d) Don't know
3. Attitude: what one says he wants or he thinks. Example
Should smoking be banned in public places? Yes, No, mixed feelings
How do you agree or disagree on banning smoking in public places?
 - a) Agree strongly
 - b) Agree
 - c) Disagree
 - d) Disagree strongly
4. Belief: what one thinks is true. Example
Do you think banning smoking in public places is good for non-smokers? Yes, No.
5. Behavior: what one does, has done or will do. Example
Do you sometimes smoke in public places? Yes, No
6. Experience: what has happened to one. Example
Have you ever had a heart attack? Yes, No.

Types of questions

Depending on the way answers are laid or obtained, questions could be divided to open-ended and closed-ended questions. Open-ended questions have no provided answers to choose from, the respondent has to create his own answer. On the contrary, closed-ended questions have a list of answers (either ordered or unordered) provided from which the respondent can choose one (or more than one) answer. Closed-ended questions sometimes contain an open item which allows the respondent to add his own answer if he could not find it from the list (partially closed-ended).

Open-ended questions are good in exploratory research when we want to obtain information for future use in questionnaires; when we want to probe the respondents

memory; when there are a large number of possible answers to the question we want to ask; and when we think that providing ready answers could bias the respondent's answer,

Closed-ended questions are better to use when we want to obtain information on simple facts for which there are few possible answers; when we want to assess the strength of feeling of the respondent; when we want the respondents to choose priorities from a list of possible choices; and when we think open-ended questions are difficult to process and analyze.

Weaknesses of open-ended questions are:

1. Burden on respondent
2. Difficulty of analysis
3. Need for recall
4. Probing is necessary
5. Not suitable for postal questionnaires
6. Illegibility of answers
7. Omissions are difficult to determine

Weaknesses of closed questions are:

1. Exclusive list of answers needed
2. Information may be missed
3. List may need to be too long
4. Long list may be burden on respondent
5. bias towards first answer

Types of questionnaires

Questionnaires can be divided to three types depending on how they are used i.e. whether they are used in face-to-face interviews (with subjects or proxies), telephone interviews or as postal questionnaires (self-administered). Selection of the questionnaire type depends on the nature of the study, the research setting and resources available for the study. Face-to-face questionnaires are used by an interviewer to obtain information from respondents directly. Telephone questionnaires are used by an interviewer talking to the respondent through the phone. Postal questionnaires are mailed to respondents where they are completed by the respondent in absence of any interviewer.

When we design and format a questionnaire, we have to remember how the questionnaire will be used. Advantages and disadvantages of questionnaire types are mentioned below.

Strengths of face-to-face interviews:

1. Better for obtaining a representative sample
2. Better response rate
3. More flexibility in questionnaire design

Weaknesses face-to-face interviews:

1. Social desirability and interviewer distortion may cause more errors
2. Need to train interviewers

Strengths of mail questionnaires:

1. Locating respondents may be easier
2. Less social desirability effect and interviewer distortion
3. Cheaper and easier administratively

Weaknesses of mail questionnaires:

1. Need to have reliable mail services and high rate of literacy
2. Not possible to determine non-respondent characteristics
3. Need to use a short questionnaire and less complex questions
4. Contamination by others
5. No control on the sequence of answering questions
6. More missed or un-answered questions

Strengths of telephone interviews:

1. Better in obtaining a representative sample
2. Easier sample selection
3. High participation rate
4. More flexibility in questionnaire design than postal interviews
5. Non-response can be minimized
6. Cost less than face-to face interviews
7. Interviewer training is easier
8. quicker than other methods

Weaknesses of telephone interviews:

1. Need to have national coverage phone service
2. Difficult to confirm respondent identity
3. Need to avoid more complex questions
4. Possibility of respondent consultation while answering

Physical layout of the questionnaire

Whether self-administered or administered by an interviewer, the questionnaire should be properly formatted and laid out to facilitate use and minimize error. The questionnaire should be attractive in terms of length, paper, printing, legibility, spacing etc. and it should be convenient for use by the interviewer or subject.

- Title: Name of institution and research/questionnaire
- Introduction to encourage participation
- Identification: A unique identification number for each questionnaire
- Paper and binding: use durable paper and different colors if possible. A booklet form of a questionnaire is better to use and can prevent loss of pages
- Size: Better to be shorter; don't put in a lot of questions which you may not use in the analysis
- Font: use a sufficiently large font convenient to read by the interviewers and subjects
- Numbering: number pages consecutively. Number questions consecutively

- Text formatting: format text nicely to help reading; adequate margins, line spaces, indentation, different font sizes(styles) to separate instructions(usually italic) questions and answers and important words in a question
- Answer lists: putting answer lists in a column (vertical) rather than in a row facilitates filling and data entry and minimizes errors
- Coding: same code for same answer e.g. always 1. Yes 2. No
- Use of illustrations, arrows when needed such as for skip patterns.
- Space for interviewer comments, respondent comments.
- Cover letter: for mail questionnaires

Proper wording of the questions

Proper wording of individual questions is an important part of a good questionnaire. If not properly worded, a question may not get understood, may elicit a wrong answer, different answers from different subjects or may be insulting or culturally inappropriate. The following points should be remembered while writing individual questions:

- The words and questions should be uniformly understood by all subjects
- No use of abbreviations, technical words and jargons
- No vague words and questions (*Have you seen a doctor recently?*)
- Not too precise questions (*During the past year, how many days have you had diarrhea?*)
- Not too general questions (*Do you sometimes feel bored?*)
- Not biased and suggestive questions (*Don't you think that hand washing can prevent diarrhea?*)
- Proper timing and wording of threatening and objectionable questions (*How much in Iraqi Dinars is your monthly salary?*)
- No more than one concept in a question (*During the past week did you have fever and cough?*)
- No use of double negatives (*Do you think that lack of sufficient tap water is not a cause for diarrhea?*)
- No ambiguous time frame (*How often do you get diarrhea?*)
- Not too demanding (like providing an answer list of 15 items)
- Is the question applicable to all respondents? If not use a skip pattern. (*What is the highest level of education you have attended?*)

Pre-testing the questionnaire

Pre-testing the questionnaire before actual use in the study is absolutely essential to be sure that it is acceptable and understood by participants and appropriate for the study. Pre-testing is an evaluation of the overall questionnaire and its individual question in relation to content, understanding, layout and acceptability before being used in the actual research. Feedback from pre-testing can improve the questionnaire and is crucial for undertaking successful data collection. When you have drafted your questionnaire after doing your literature search and homework, you can take the following steps to finalize your questionnaire while in each step taking the feedback and incorporating feedback to revise the questionnaire.

1. Share with some colleagues in the field of study
2. Test with some friends/family members
3. Revise
4. Train interviewers for piloting on study participants
5. Pre-test with some study subjects
6. Revise and finalize

Some of the issues we have to look for in the pre-test include the following:

- Is the questionnaire layout attractive for the interviewer and the respondent
- Is the cover letter acceptable?
- Are the instructions clear?
- Is each question measuring what it is intended for?
- Are questions interpreted similarly by all respondents?
- Are all words understood?
- Are there any biased questions?
- Are there any threatening, objectionable questions?
- Does the answer list contain exhaustive items?
- Are there any formatting/font/indentation/numbering changes that need to be made?

Validating the questionnaire

Validity and reliability of the questionnaire are important to check in order to make sure that it is an appropriate tool for measurement. Validity of a question means the extent to which that question measures the true value of the variable that we are testing. Reliability of a question measures how consistent the question is when asked by different observers or by the same observer repeatedly. Testing validity of a questionnaire is a difficult task because usually there is no gold standard to compare the findings. If gold standards existed we could determine validity by measuring sensitivity (proportion of true positives) and specificity (proportion of true negatives) for each question. Sometimes other objective measures exist with which we can compare the responses. For example reported age could be validated against the subject's ID card. Validity of continuous variables could be checked by calculating Pearson's correlation coefficient with the gold standard.

Reliability of the question means that the same answer is obtained if the question is asked repeatedly provided the circumstances have not changed. Reliability therefore can be tested by administering the same questionnaire by two observers (inter-observer reliability) or by administering the same questionnaire to the same person on 2 occasions (intra-observer or test-retest reliability). The responses to categorical variables are then compared by mean pair agreement calculating kappa statistic. Reliability of continuous variables is tested by intra-class correlation coefficient.

Further reading

- Field trials of health interventions in developing countries, edited by P Smith and R marrow
- Principles of exposure measurement in epidemiology by BK Armstrong et al.
- Guide to questionnaire construction and question writing by