

Questionnaire designing and validation

Abdul Momin Kazi, Wardah Khalid

Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan.

Corresponding Author: Abdul Momin Kazi. Email: momin.kazi@aku.edu.

Introduction and Objectives:

Measurement is an essential activity of medical science. In order to acquire data about people, objects, and events proper data collection tools need to be designed which can measure things of scientific interest. Questionnaire is one of the most important techniques to collect data. It is a common experience that many studies focus on similar objective and theme but have entirely different questionnaires. It is extremely important for a researcher to know the importance of a proper questionnaire and whether it measures what it is intended to measure. Composing of a questionnaire is always much more complex than expected and great attention is required to its flow, format and length. Making an individual question is a tedious task and validating this questionnaire is another challenge which at times is overlooked. Consideration should be given accordingly on whether the questionnaire will measure quantitative or qualitative data, and what would be its mode of administration.

The aim of this paper is to give a basic introduction and broad overview to the readers on questionnaires for epidemiological studies.

Questionnaire Development:

A questionnaire is a tool to obtain information from respondents. It is an efficient way to collect data, the investigator should be certain on what is required and how to measure the variables of interest. As a first step it is always recommended to do a literature search on previously used validated questionnaires that can be administered in similar settings and capture variables that are of interest according to the study hypothesis. These questionnaires do not need to be tested for reliability and results can be compared for different studies and also combined for meta-analysis. However one needs to make sure that the mode of administration should be similar to the original questionnaire.¹

If a new questionnaire is to be developed, it should be pilot tested and validated in order to evaluate if it is measuring what it is supposed to measure and is it doing it reliably. Wording of questions is very critical and should take into consideration; appropriateness of the content, level of sophistication of language, type and form, sequence and how is data sought from the respondents. During questionnaire development its mode of

administration should be kept in mind, whether it will be self-administered or interview based and its design and flow should be planned accordingly.² See Table on common problems in questionnaire.

The language of questionnaires should be at the level of understanding of the participants. It is essential to word the questions in a way that they can easily be understood by participant and should be according to their educational level and culture. If the questions are interpreted differently by the participants it will result in wrong answers and responses will thus be biased. Reading ease of a questionnaire can be assessed by Flesch reading ease score.³

Translation of a questionnaire is essential if an instrument is not available in a language understood by the target population. The initial stage of translation is a source language questionnaire, from which translation in required language is done. Translation is not a mechanical work and should not be done on word to word bases across languages. It is important to understand the local context, specific issues and cultural meanings which language carries. Translation should not only be concerned with translating meanings, but it should also understand how the language is tied to local realities and literary forms. The term back translation is highly recommended in questionnaires related to health surveys. Back translation helps in evaluating the quality of the translation. The source language is translated in another language and again translated back into the source language. Translation back to the source language is done by another translator who is unaware of the source language version.⁴

Validity:

Validity is the degree to which an assessment measures what it is supposed to measure. Validity is a complex topic and it is beyond the scope of this paper to explain it in detail, however to make the readers acquainted we have described it briefly. Essentially there are three types of validity i) content validity, ii) criterion-related validity, and iii) construct validity.⁵ A questionnaire undergoes a validation procedure to make sure that it accurately measures what it aims to do, regardless of the responder. Valid questionnaire helps to collect better quality data with high comparability which reduces the effort and increase the credibility of data. A valid questionnaires must have

Table: Some Common Problems in Questionnaire.

Problem	Example	Issue	Solution (Remarks)
Unclear	"Have you ever smoked"	What does smoked means? Does it only refer to cigarette, pipe and cigars as well	"Have you ever smoked as much as one cigarette a week for as long as six months"
Non self-explanatory	"How heat is delivered to your home"	The investigator is interested in whether heat is delivered in form of hot air, water or steam. However question maybe answer as "heat was transferred in fuel truck"	All possible answers should be written or read immediate after the question
Two Questions combined	"Do you get out of breath when you jog"	Here "no" can refer to no shortness of breath or doesn't go for jogging	"Do you go for jogging" and those who answer yes should be asked regarding breathless ness
Difficult words to understand	"What was your age of menarche"	Many female may not understand the word menarche	"How old were you when your menstrual cycle started"
Events difficult to remember	"How many times a day you used to eat apple when you were between the age of 8 and 10 years"	Researchers often take for granted that people will remember accurately things from the past	Questions related to events that occurred long time back should be measured with caution
Hypothical Questions	"Would you like to go for vacations"	If money permits most of us would like to take vacations	Certain situation correlate poorly with actual situation
Categories	"When you go camping do you sleep in a camper, trailer or tent "	Some people might sleep in the open as well	Option for "somewhere else" should have been asked
Too long question	"Have you ever used antibiotic, anti-malarial, tranquilizers, or diuretics"	The question is straight forward but it is too long. It is human nature to remember the last item only	This could be divided in different questions and medicines taken should be asked separately
Question details	"Tell me each brand of medicines you have used and for how long since you last got pregnant"	Respondents are being asked too many questions at once	This should be asked in systematic series
Leading Question	"Do you do your exercises regularly"	High chance that respondent will answer yes for exercise.	Do you do your exercises for at least 20 minutes everyday.

following characteristics (i) simplicity and viability (ii) reliability and precision in the words (iii) adequate for the problem intended to measure (iv) reflect underlying theory or concept to be measured and (v) capable of measuring change.⁶

Type of Questions:

A questionnaire is a written document to gather information irrespective of mode of administration. A questionnaire could be structured, in which all the participants are asked same questions in the same way, this is usually interview based questionnaire format. The other type is unstructured questionnaire and questions may vary at the discretion of the interviewer. Unstructured format may be used at clinical setting however structured questionnaire is preferred for epidemiological studies as same data from all respondents need to be analyzed and measured.⁷

Open-ended questions allow respondents to answer them in any way they want. For continuous variable this format is more suitable when large numbers of options are available and it is not practical to write all answers in advance e.g. weight of the patient. All possible answers are not written in advance

and needs to be coded latter and may increase chance of error. The open ended questions might increase the burden on work and responses have to individually review by the investigator before assigning codes and analyses.

Closed-ended questions in contrast would ask the respondents to make choices among a set of answers in a given question. The response could be mutually exclusive or may select more than one option. For measuring dichotomous variables closed- ended questions are preferred because possible answers can be easily precoded. Precoded questions are defined as those in which numbers are assigned to a given answer. Precoding saves time for assigning number latter and hence decrease error; however for open-ended questions coding is done after the data is collected. Coding helps in data entry, as information of questionnaires in paper format are entered in data entry programs by putting in the numbers rather than writing the whole answer.

Questionnaire Style and Appearance:

The appearance and style of the questionnaire is very important and has a very strong impact especially in self-

administered questionnaire. Format, order, spacing, fonts used and grouping of the response are very important features in overall layout of the questionnaire and have a direct effect on the responses and time spent by the respondent to provide it. Questions should be simple, clear and easy to understand, using minimum of words and space and only asks what needs to be asked. Lengthy or confusing lay out of the questionnaire can also make the interviewer confused and responses administered by the interviewers may not be accurate or complete. The clarity of questionnaire has direct impact on data collected by the interviewer and responses given by the responders.

Mode of Administration:

Broadly speaking there are two modes of administering a questionnaire, a) self-administered and b) interviewer administered questionnaire.

Self-administered questionnaire only requires questionnaire distribution; it is much cheaper and doesn't need trained staff. This mode is less susceptible to information bias and interviewer effect but have greater chance of having no response items. The main advantages of self-administered questionnaires is that it can reach a large sample size, cover wide geographical area, cover population which is sometimes difficult to reach, excellent for capturing sensitive topics and cheaper as compared to other modes of administration.

Common method of self-administered questionnaire distribution is either through mail or electronic distribution. Participants can complete mailed questionnaire at their convenience, in their homes and at their own pace. Major disadvantage of the mailed questionnaire is low response rate even after repeated mails, and queries of the participant cannot be clarified. Some effective techniques for improving the rate of response are; sending follow-up letters, enclosing some incentives, providing self-addressed stamped envelope and keeping the questionnaire brief.

Electronic and web-based questionnaire, including data collected through personal digital apparatus (PDAs), smart phones and cell phones are latest techniques for questionnaire administration. Questionnaire can be designed to filter and screen participant's response, checks for input error, range and skip patterns can be incorporated preventing significant typing and data format error.⁸ However electronic questionnaire is restricted to those participants who have access to a computer and internet and this can be a potential bias.

In person or interview based administration is expensive but provides direct interaction with the participant. The interviewer has the opportunity to introduce the research topic and motivate the participant to offer their frank answers and questions can be clarified at the spot. If the interviewer is trained and motivated, it is the best method to collect data in epidemiological studies. In recent era telephone use has also increased for administering questionnaire. This is cost effective and usually have greater response rate as compared to postal questionnaire. Telephone interviewing also facilitates in covering a large number of participants over a wide geographical area.⁹ Interviewer on the phone can directly talk to the participant, explain the study and clear any confusion or questions during the call. Possible disadvantage of phone call is that the caller cannot see the participant and might have difficulty in rapport, cannot use additional material for explaining the questionnaire. For example pictures and study involving sample collection cannot be conducted through phone administered questionnaires.

Conclusion

A questionnaire designed for epidemiological studies should capture information from participants regarding their exposure, possible risk factors, and occurrence of disease of interest. This paper tried to help researcher in designing questionnaire and broadly explained the different method of questionnaire development, validity, types of questionnaire, their style and appearance and mode of administration.

Reference

1. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010; 11: 2.
2. Streiner DL NG. Health measurement scales: a practical guide to their development and use: Oxford University Press, 2004.
3. Farr JN, Jenkins JJ. Tables for use with the Flesch readability formulas. *J Appl Psychol* 1949; 33: 275-8.
4. Harkness J. Questionnaire Translation. In J.A. Harkness, F.J.R. van de Vijver, and P.Ph. Mohler. *Crosscultural Survey Methods* New York: John Wiley, 2003.
5. Devellis RF. *Scale Development* SAGE Publications; 2003.
6. Garcia de Yebenes Prous MA, Rodriguez Salvanes F, Carmona Ortells L. [Validation of questionnaires]. *Reumatol Clin* 2009; 5: 171-7.
7. Kelsey JL. *Methods in observational epidemiology*. New York: Oxford University Press, 1996.
8. Seebregts CJ, Zwarenstein M, Mathews C, Fairall L, Flisher AJ, Seebregts C, et al. Handheld computers for survey and trial data collection in resource-poor settings: development and evaluation of PDACT, a Palm Pilot interviewing system. *Int J Med Inform* 2009; 78: 721-31.
9. Potthoff RF. Telephone sampling in epidemiologic research: to reap the benefits, avoid the pitfalls. *Am J Epidemiol* 1994; 139: 967-78.