**Title of the article:** How to design and validate a questionnaire: A guide

**Abstract Page**

**Contributors**

1 Jaspreet KaurSidhu, MD, 2 SurjitSingh, MD, DM, 3Ajay Prakash, PhD, 4Bikash Medhi, M.D. 5Preeti Sidhu, MBBS.

**Department(s) and institution(s)**

1Department of Pharmacology, Gian Sagar Medical College and Hospital, Patiala, Punjab, India-140601.Email id: jaspreetkaurboparai@gmail.com

2Assistant Professor, Department of Pharmacology, AIIMS, Jodhpur, Rajasthan – 342005 Email id: sehmby\_ss@yahoo.com

3Assistant Professor, Department of Pharmacology, Post Graduate Institute of Medical Education & Research, Chandigarh. E mail id: ajayprakashpgi@gmail.com

4Professor, Department of Pharmacology, Post Graduate Institute of Medical Education & Research, Chandigarh. E mail id: drbikashmedhi@gmail.com

5Junior Resident, AIIMS Jodhpur, Rajasthan. Email id: preetisidhu72@gmail.com

**Abstract:** A questionnaire is a commonly used data collection method and is a very crucial part of research. However, designing a questionnaire can be a daunting task for postgraduate students. This manuscript illustrates the various steps required in questionnaire designing and provides an insight on the essentials of questionnaire construction and validation. Data from a questionnaire should be able to comprehend the objectives of the study; else it may lead to wrong interpretation or bias, decreased power of study and inability to generalize the study results. Since it is equally important to verify the usefulness of the designed questionnaire, the article briefly describes the process of psychometric evaluation of a questionnaire.

**Key-words:** Questionnaire validation, Questionnaire designing, Psychometric assessment, Validity.

**Key Messages:** This will help the postgraduate students in better understanding of designing and validating the questionnaire for their work.

**Introduction:**

Text

A questionnaire as defined by Babbie*“is a document containing questions and other types of items designed to solicit information appropriate to analysis”*.1It is a cost effective research instrument to assess relevant information for the study. Getting the answers to a research question with the help of a questionnaire is a step-wise process. Once you have a research question in mind, it is advised to search for validated questionnaires employed in studies with identical variables of interest for the present study.2 Using a previously published questionnaire saves time and resources and offers the advantage to compare your results with other studies which may be incorporated in a meta-analysis and is also easier to publish. However, it has to be administered in the same way as the original questionnaire after obtaining permission from the authors or publishers.3 Validated questionnaires by World Health Organization or by various recognized disease societies and organizations can be used. For example, to assess pain, stiffness & physical function in patients with knee osteoarthritis, a previously validated scale by the Western Ontario & Mc Master Universities (WOMAC) can be used after getting permission.

**How to Design a new questionnaire:**

A new questionnaire needs to be developed if no previously used standard questionnaire is available or if the existing questionnaire has certain limiting factors for the purpose of this study. Several points have to be considered while designing a new questionnaire including assessment of data you need, item selection for inclusion, designing the questions, use of appropriate wording for comprehension, layout and presentation of questions, coding, preparing and pre-test the first draft, piloting, evaluate the form, if needed applying pre-test again and lastly, performing the survey. The various steps involved in questionnaire designing have been summed up in Figure 1.

**Specify the study objectives**

Before the study is started, the researcher and co-researchers must be clear on what information they intend to generate. A formal discussion regarding the goals of the study, the target population, length and format of the questionnaire and even the mode of administration is very beneficial to come to a consensus regarding the study design and to have clear objectives and hypothesis.4

**Framing the questions**

Once a clear hypothesis is there, questions can be framed accordingly. It is important to have clarity on the importance of each question in relation to the objective of the study. It is suggested to list variables or indicators related to the objectives and hypothesis of the study before designing a questionnaire.5For example, an effective questionnaire on anxiety would have questions on various variables like autonomic symptoms, depressed mood, cardiovascular symptoms etc.

**Language**

The language of the questions should be as per the education level and culture of the study group and should be easily understood by the participants.2 For example, are you a primigravida might not be understood by many females. The wording of the questions should be such that all participants interpret the same meaning of the question.

**Types of question**

Based on the possible answers to a question, they can either be open-ended or closed ended. Open-ended questions are used when a large number of possible answers exist and the researcher cannot predict all the possible answers beforehand.6 For example, name 5 commonly used herbs in India? A space is provided for the participants to write the text for the answer. The advantages of both open and closed ended questions are given in Table 1.

However, each response option for the open-ended questions has to be inspected before assigning codes. Coding and analyses of this type of questions is at times a time-consuming process.7 Another disadvantage is that sufficient amount of space is needed for the answers, which makes the questionnaire look lengthy and have a lower response rate.

Closed-ended questions require the participants to select the answer from a set of choices for a given question. For example, circle your source of information about new drugs? i) Medical journals, ii) Internet, iii) Books. Numbers can be assigned to a given answer to pre- code the questions. However, a disadvantage of missing an important answer option is there with closed-ended questions.6

The way questions are framed is important and a checklist for constructing good questions is given in Table 2.

The questions should preferably be arranged in the following manner- i) general followed by specific, ii) easy followed by difficult, iii) concrete to conceptual, iv) closed ended followed by open-ended, v) demographic and personal questions in the end.8 It is also recommended to provide short and simple instructions to attempt the questions. E.g. Mark the correct response.

Attention should be given to the overall look, format and clarity of the questionnaire as these significantly affect the response rate and thus have an impact on the success of the research project. A check on the font-size, color, spacing, underlining, bold type is advised to get a good graphically designed questionnaire.3E.g. Increase the font size for the elderly participants.

**Piloting the new questionnaire**

For a small survey pretesting of the drafted questionnaire as a whole is sufficient. In case of a large survey, piloting is done in three phases. The first phase involves taking details from each participant about a few questions. In the second phase, the entire questionnaire is read out to the respondents by interviewers. The questionnaire is improved based on interviewer`s comments and scrutiny of the responses of the participants. The third phase involves further refining of the questionnaire to improve the layout and appearance, question sequence and edit questions.8 Questionnaire should be initially evaluated by experienced researchers in the same field, followed by a sample of non-expert group which includes friends, relatives or colleagues.3 To understand how participants interpret questions and attempt the answers ‘cognitive interviewing’ can be used.9

**Mode of administration**

A questionnaire can either be self-administered or read out by interviewers. Self-administered questionnaires only require questionnaire distribution, thus they are cheaper and more convenient to administer. No trained staff or interviewer is needed as the questionnaire may be sent by post, email or electronically online. These questionnaires have least investigator participation and are less prone to information bias and interviewer effects. The main advantages of a self-administered questionnaire are that it can cover more geographical areas, larger sample size and also population which may be difficult to reach and are excellent for covering sensitive topics. However, they are more susceptible to item non-response.

Interview or in person administration is expensive and time-consuming but allows direct interaction with the participant. The interviewer gets to develop a rapport with the participant and brief him about the research study. It not only allows interviewer to clarify questions but also to check the answers.2, 3

**Psychometric assessment of a questionnaire:**

A well-designed questionnaire generates good quality data with high comparability and credibility and must undergo psychometric assessment. The psychometric assessment of a questionnaire involves evaluation of four main domains naming feasibility, reliability, validity and sensitivity to change.10Reliability and validity are however the most important concerns of a newly designed questionnaire.

**Feasibility**

The feasibility of a questionnaire is determined by a pilot study in approximately 30 subjects. The points to consider while determining feasibility include time taken to fill the questionnaire, simplicity of the format, clarity of the questions, ease of scoring and result interpretation.10The oldest and the most precise formula to assess reading ease is Flesch reading ease score. A score between 60-69 is considered acceptable.11

**Reliability**

Reliability is an indicator of precision of the questionnaire i.e. it evaluates if the questionnaire consistently determines what it intends to determine. A low reliability demonstrates that the results of the questionnaire may vary greatly if the data is collected again with the same participants. For example, if a participant scored 10 out of 50 on a HIV knowledge questionnaire and scored 45 the next time, the questionnaire becomes unreliable. It is difficult to assess if the knowledge of the participant about HIV was high (45) or low (10). Reliability is determined through a pilot test on a small sample of participants. Reliability is measured in aspects of stability, internal consistency and equivalence. Stability is measured by test-retest method, internal consistency by split-half method and Cronbach`s α statistics and equivalence by inter-observer reliability. Test-retest reliability is a commonly used method and represents the degree of association between the test and retest scores of the same group of individuals taken about one week apart.12Correlation coefficient is of test-retest scores is calculated and analysed.

In split-half method, questionnaire contents are split into two equivalent halves, either first-second half or odd-even numbers and then the analysis ofcorrelation coefficient of scores of two halves is carried out. Correlation coefficient of ± 0.7 to 1.0, ± 0.3 to 0.69 and ±0.0 to 0.29 is regarded as strong, moderate and weak correlations respectively.Statistical techniques like Cronbach`s α statistics or item-total correlationscan be used to demonstrate internal consistency. It uses inter-item correlations to show if the consisting questions are measuring the same domain. Cronbach’s alpha coefficient is calculated for total score and then change is observed in coefficient eliminating one item at a time and items that change the alpha coefficient significantly should be re-revaluated and modified. Cronbach`s α should be more than 0.70 for each domain of the questionnaire.13Equivalence is determined by inter-observor reliability which measures agreement between observes and it is accomplished by multiple researchers scoring the same instrument and determining percentage of agreement between them.

**Validity**

A questionnaire can be reliable but invalid but a valid questionnaire is always reliable. Validity is the degree to which a questionnaire evaluates what it is intended to evaluate. It is of four types face validity, content validity, construct validity and criterion validity. The type of validity evaluation depends on the goals of the study.14It evaluates the amount of built-in error in the questionnaire.12 For example, a mail questionnaire used to assess the level of text anxiety in school going children written at the level of a college going student is an invalid questionnaire. It is important to get answers to the following questions while looking for validity in your questionnaire- does it actually measure what it intends to, is it suitable for the designed study and target population, does it collect all the information pertaining to the objectives of the study?14

Face validity refers to the degree to which a questionnaire “appears” to evaluate what it was designed to evaluate. It is determined by assessing the comments of experts in the clinical area and the study participants on the questionnaire.12 The missing data should not be more than 5%. Variance in extremes is indicated by percentage of patients with the minimum and maximum points and it should not exceed 20%.15

Content validity refers to the extent to which the questionnaire describes most of the dimensions of the concept under study. It is important to envelop all related aspects of the concept under study for the questionnaire to be deemed valid. The judgmentsare based on expert opinion, review of medical literature, results of new pilot studies all of which can aid in the evaluation of content validity. Factorial analysis can also be used for the evaluation of content validity. The answers to the questions are explored and grouped in relation to the underlying factors that identify the possible dimensions.10Average Congruency Percentage (ACP) designed by Popham (1978),16 compute the average from the percentage of questions deemed to be relevant to the experts with regard to content. The value greater than 90 ensures content validity.Content validity index (CVI), conceptualize by Martuza16 for individual items (I-CVI) and for whole scale (S-CVI) can also be used for content validation. A minimum of threeand maximum of ten experts asked to review the relevance of each question on a 4-point Likert scale with one score as non-relevant and four as most relevant. Content is validated if more than 80% (I-CVI = 0.80) of experts think that the content is relevant. Eg: We have a scale consisting of ten items. If for a single item, 8 out of 10 experts give score greater than 3, then I-CVI is 8/10 = 0.80 for that item (Table 3). For scale validation i.e. for S-CVI, S-CVI/UA which is defined as the proportion of items of a scale that has been ascribed a relevance rating of three or four by all the experts and S-CVI/Ave which is average of the I-CVIgiven by experts for all the items on the scale, can be estimated. Eg: We have a scale of ten item. If only five items have been given a relevance of three or four by all the experts, then S-CVI/UA will be 0.5 and rest five items have I-CVI of 0.80 meaning that 80 percent of experts have given a relevacne score of three or four (Table 3). The S-CVI/Ave will be 0.9 calculated by taking the average of the I-CVI for all the individual items. The best method to ensure stringent validity is S-CVI/UA but it is difficult to use if multiple experts are validating. In such situtations, we can prefer S-CVI/Ave.

Construct validity is the assessment of the degree to which the questionnaire reflects the theory of the phenomenon or concept of the designed study or relationship between questionnaire and underlying theory. It`s a continuing process, wherein the conduct of the questionnaire is analyzed in many studies done in varied settings and participants.17E.g. SF-36 is a health-related quality of life questionnaire used extensively in about 4000 publications, reflecting that scores of SF-36 will reflect the clinically and quality of life disease outcome parameters.

Criteria validity generally describes how well the questionnaire correlates with anexisting 'gold standard' measure that can predict the disease outcome and is taken as a reference to evaluate validity. Such gold standards may be clinical assessment or clinical investigations or other validated questionnaires. Determination of criteria validity involves identification of relevant and reliable external criteria, reunion of a sample of population in which the questionnaire will be used, administration of the questionnaire and scoring of results and finally comparing each individual with the external reference criteria.10Eg: The results of a questionnaire assessing disease severity of angina patients may be compared with the results of their Coronary angiogram.

Sensitivity to changeis the ability of a scale to detect a clinically significant change over time in the same individuals or different group of participants.10

**Translation**

Translation is needed in case a previously used questionnaire is not available in the language understood by the researcher or the study population or if a researcher plans to conduct a multi-centric trial using a newly developed questionnaire across different culture groups and also if the country of intended use has high percentage of immigrants. The initial translation from a source language to the required language is followed by back translation into the source language which is useful in determining the quality of translation. To ascertain that the questionnaire is understood in the same way as the original questionnaire and there is no loss of information, it is recommended to do pilot testing after translation has been done.2, 7Article by Fukuhara et al provides an insightful reading on the translation and validation of SF-36 health survey into Japanese language.18

**Box 1: Advantages of open and closed ended questions**

**Open ended questions**

* Allows exploration of range of possible answers

**Closed ended questions**

* Easy to fill
* Less time needed
* Less discrimination based on literacy level of participants
* Easy to code ,score and analyse
* Easy to report results
* Increased response rates

**Conclusion:**

A questionnaire is a quick method of data collection critical for pragmatic research. A poorly designed questionnaire generates poor quality data which is difficult to interpret. This article presents a framework for effective questionnaire development and emphasizes the need to adopt a logical and structured approach to questionnaire development. It also highlights strategies to demonstrate the validity of a newly developed questionnaire. There is a need to plan each stage of questionnaire development and consider its psychometric evaluation to avoid collecting misleading information which results in evasive conclusions and futile recommendations.

**Acknowledgements:** None

Table 1: Advantages of Open and Close ended questions

|  |  |
| --- | --- |
| Open ended questions | Close ended questions |
| Relatively easy to frame | Greater response rate |
| Possible to analyse variety of answers | Easier to code and interpret results |
|  | Less bias based on literacy level of participants |

Table 2: Checklist while framing questions

|  |
| --- |
| 1. Short and clear |
| 1. Relevant to the study |
| 1. Simple language |
| 1. One question at a time |
| 1. Appropriate response options |
| 1. Avoid leading questions |
| 1. Avoid technical terms and slang |
| 1. Avoid reference of previous questions |
| 1. Avoid embarrassing or sensitive questions |
| 1. Avoid calculations and abbreviations |
| 1. Minimal negative questions |
| 1. Questions on the same topic is same section |

Table 3: Calculation of I-CVI, S-CVI/UA and S-CVI/Ave by fictitious rating and Content validation by Martuza (1977).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Experts | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Number in Agreement | Item CVI |
| Item |  |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0\* | 0 | 1\* | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 8 | 0.8 |
| 2 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 8 | 0.8 |
| 3 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 8 | 0.8 |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 8 | 0.8 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 8 | 0.8 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 1 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 1 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 1 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Expert Proportion | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 |  | Mean I-CVI = 0.9; S-CVI/UA = 0.50;  Mean Expert Proportion = 0.9 |

I-CVI - Individual Item - Content validity index; S-CVI/UA – Scale - Content validity index/Universal;

S-CVI/Ave - Scale - Content validity index /Average;

Item rated three or four on four point relevance Likert scale. 0 for response 1or 2; 1 for response 3 or 4.

Figure Legends:

Figure 1: Steps involved in conceptualization, designing and establishing validity and reliability of Questionnaire

**References:**

1. Babbie E, editor. Survey research methods. 2 ed. Belmont: Wadsworth Publishing; 1990.
2. Abdul Momin Kazi WK. Questionnaire designing and validation. J Pak Med Assoc 2012; 62: 514-6.
3. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. Trials 2010; 11:2.
4. Boynton PM, Greenhalgh T. Selecting, designing, and developing your questionnaire. BMJ 2004; 328: 1312-5.
5. Acharya B. Questionnaire Design June 2010: Available from: <http://www.saciwaters.org/CB/IFRM/IFRM/IV.%20Literature/Module%206_Qualitative%20Research%20Methods/6.4%20Questionnaire%20Design_Acharya%20Bidhan.pdf>.
6. Brancato SMG, Murgia M, Signore M, Simeoni GKB, Körner T, Nimmergut APL, et al. Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System: European Commission Grant Agreement 200410300002; 2006. Available from: <http://epp.eurostat.ec.europa.eu/portal/page/portal/research_methodology/documents/Handbook_questionnaire_development_2006.pdf>.
7. Nieuwenhuijsen MJ. Design of exposure questionnaires for epidemiological studies. Occup Environ Med 2005; 62: 272-80.
8. Leung WC. How to design a questionnaire. STUDENT BMJ. [Education]. 2001; 9: 187-9.
9. Willis GB. Cognitive Interviewing A “How To” Guide: Available from: <http://appliedresearch.cancer.gov/archive/cognitive/interview.pdf>.
10. Garcia de Yebenes Prous MA, Rodriguez Salvanes F, Carmona Ortells L. [Validation of questionnaires]. Reumatol Clin 2009; 5: 171-7.
11. Flesch R. The Flesch Reading Ease Readability Formula. [cited 2013 29 October]; Available from: <http://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php>.
12. Norland EVT. Controlling Error in Evaluation Instruments. Journal of Extension [serial on the Internet]. Summer 1990; 28: Available from: <http://www.joe.org/joe/1990summer/tt2.php>.
13. Rattray J, Jones MC. Essential elements of questionnaire design and development. J Clin Nurs 2007;16: 234-43.
14. Radhakrishna RB. Tips for Developing and Testing Questionnaires/Instruments. Journal of Extension [serial on the Internet]. February 2007; 45: Available from: <http://www.joe.org/joe/2007february/tt2.php>.
15. El Moussaoui R, Opmeer BC, Bossuyt PM, Speelman P, de Borgie CA, Prins JM. Development and validation of a short questionnaire in community-acquired pneumonia. Thorax 2004; 59: 591-5.
16. Polit DF and Beck CT. The Content Validity Index: Are You Sure You Know What’s Being Reported? Critique and Recommendations. Research in Nursing & Health, 2006, 29, 489–497.
17. International Consultation on Incontinence Modular Questionnaire (ICIQ). [cited 2013 29 October]; Available from: <http://www.iciq.net/validationprotocol.htm>.
18. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol 1998; 51: 1037-44.