

Questionnaire in Clinical Research and Practice

T O Lim, FRCP, Clinical Research Centre, Kuala Lumpur Hospital, Jalan Pahang 50586 Kuala Lumpur

Measurement is the means of collecting data from patients whether for purpose of research or for decision making in clinical practice. The commonest method of measurement used in research or practice is simply by observation. An obvious example is the physical examination of patients by a clinical investigator or by a doctor. Another familiar method is the use of equipment like in chemical measurement of serum creatinine or physical measurement of a patient's body weight. In this issue of MJM, there are 4 papers¹⁻⁴ that utilize another type of measurement instrument, the questionnaire.

Questionnaire is a means of collecting data from patients where they provide verbal or written responses to a set of questions. Thus, Quek *et al* uses the Beck Depression Inventory and HDQOL-20 to measure depression in and quality of life of urological patients respectively^{1,2}. Sararaks *et al* uses the SF-36 to measure the quality of life of asthmatics³ and Ho *et al* uses Griffiths Mental Scales to evaluate the development of children⁴ (the Griffith Scales is only in part a questionnaire). What are MJM readers to make of these measurements that are based in part or entirely on patient's subjective responses? Are they likely to be as "good" as measurements taken with a well-calibrated equipment or observation by an expert clinician?

For a start, the use of questionnaire is often unavoidable. The only way to measure abstract attributes like quality of life, depression, mental or linguistic development is by asking the patients directly. No equipment exists to measure these attributes nor are they directly

observable. And we should not automatically suspect the quality of the data just because the measurement is relatively subjective. Questionnaire instruments can be made as rigorous as any equipment or observational method if properly constructed and validated.

In reading any research based on data collected by means of questionnaire, the reader should ask himself or herself two questions.

1. What evidence is there that the questionnaire has been successfully adapted to the culture of the population on which it is applied?
2. What evidence is there that the questionnaire is reliable and valid in the population?

Cross-cultural adaptation of questionnaire

Almost all questionnaires used in research in this country were originally written in English and intended for English speaking subjects. Not surprisingly, the choice of words and idioms in the questionnaire, its syntax, and the contents or experiences it refers to are inevitably culture bound. Hence, before a questionnaire can be used in a research, we expect to see evidence that the questionnaire originally developed in another culture has been successfully cross-culturally adapted to this country. This requires translating and validating the questionnaire such that the translated version is "culturally equivalent" to the original questionnaire. Unfortunately this is an extremely complex, not to mention expensive process. Indeed, cross-cultural adaptation of existing questionnaire is a research in its own right. Perhaps understandably, almost all

researchers prefer to take the short cut; they simply assume the questionnaire, roughly translated into one or more of the local languages, is "equivalent" to the original instrument.

Reliability and Validity of questionnaire

Assuming we have a "culturally equivalent" questionnaire in hand, the next step is to evaluate its measurement properties. These are characteristics that determine the quality of the measurement. Technically, these are the types (random and systematic) and amount of error in the measurement. On the basis of the evaluation of these errors in the measurement, we determine the acceptability of the method of measurement whether for research or for clinical practice. For example, no clinicians would ever accept, say a new chemical measurement for serum creatinine, without knowing its repeatability and reproducibility (would you accept a method with coefficient of variation (CV) of 25%?), as well as its accuracy (serum creatinine is a notoriously inaccurate measure of glomerular filtration rate). Similarly, for a questionnaire that claims to measure quality of life or depression or anything at all, we expect to know its repeatability and reproducibility (read reliability for questionnaire) and accuracy (read validity). Unfortunately, like for cross-cultural adaptation of questionnaire, the process of determining the measurement properties (reliability and validity) of a

questionnaire is also rather complicated, and is usually undertaken as a research in its own right. For the same reason, almost all researchers in this country conveniently choose to assume the questionnaire they used on their subjects has the same properties as originally designed by the developer. Quek et al in this issue of MJM should be commended for undertaking the research to determine the reliability of Beck Depression Inventory and HDQOL-20 in urological patients^{1,2}. They found for example that the HRQOL has a test-retest intraclass correlation coefficient (ICC) of 0.91, which is excellent. ICC is a statistical measure of reliability (akin to CV for laboratory test), and we generally expect a research instrument to have an ICC of at least 0.7. They also determine the responsiveness or sensitivity to change of HRQOL-20, which did rather well in this respect. Responsiveness is the power of a measurement to detect a change within subject when a significant change had occurred (akin to critical difference for laboratory test). However, in my opinion, the determination of the validity of HRQOL-20 is less than satisfactory. Nevertheless, this is a nice example of the sort of research evidence we expect to have before we can accept a method of measurement for application in research or in practice. And it does not matter whether the method is a questionnaire or a fancy equipment or a high tech bioassay, the same principles and rules ought to apply though all too often ignored.

References

1. Quek *et al.* Beck Depression Inventory(BDD): a reliability and validity test in the Malaysian urological population. *Med J Malaysia*, 2001; 56: 285-92.
2. Quek *et al.* A Malaysian study on the reliability and validity of the health related quality of life questionnaire (HDQOL-20) in urological patients. *Med J Malaysia*, 2001; 56: 293-301.
3. Ho JJ *et al.* Use of Griffiths Mental Scales in normal 2 year old Malaysian children. *Med J Malaysia*, 2001; 56: 331-35.
4. Sararaks S *et al.* Quality of Life - How do Malaysian asthmatics fare? *Med J Malaysia*, 2001; 56: 350-58.