



EDITORIAL

Validated questionnaires should not be modified

E.F. Juniper

The reformatting and changes that were made to the Asthma Control Questionnaire (ACQ) [1] for the Post-cold Asthma Control and Exacerbation (PAX) study [2] raise some important concerns about modifying validated questionnaires.

Just like mechanical and electrical measurement instruments, validated questionnaires are precision measurement instruments. The only difference being that they measure subjective rather than objective health status. Like any mechanical or electrical instrument, a great deal of care and expertise goes into the development of these questionnaires. Many studies have provided developers with the knowledge of how to: specify what the questionnaire is intended to measure (its construct); structure and select the right questions; formulate the responses; select the time specifications; optimise the page formulation for accurate completion; conduct validation studies (measurement properties and whether the instrument is measuring what it is meant to measure) and provide users the wherewithal to place a clinical interpretation on the data. For the same reason that one would never think of changing the numbers on the dial of a mechanical spirometer, one should never change a validated questionnaire. Even very small changes can destroy its validity.

QUESTIONS

Questions are selected by well-established methods (usually either “importance” or “factor analysis”) [3] and their position in the questionnaire carefully ordered. Wording is checked for ease and accuracy of understanding (cognitive debriefing). For the analysis, each question has a weighting. For some questionnaires, this means that an algorithm must be used (*e.g.* The Short Form (SF)-36 Health Survey) [4], in others (*e.g.* the ACQ) questions are selected in such a way that they have equal weighting and the overall score is the mean of all the responses. The wording of questions should never be changed. Shortened versions should only be used when they have been validated [5]. Questions should never be added even if they are not going to be included in the analysis because they may alter how patients respond to other questions in the questionnaire.

RESPONSES

Most medical questionnaires use Likert-type scales (*e.g.* a seven-point scale ranging between 0=no impairment and 6=severe impairment). These are interval scales with equal spacing between each response. The verbal descriptor for each

number is carefully chosen to enhance the perception of equal distance between the numbers. Patients use both the numbers and the words to select their responses (varying in preference between patients) and so equal prominence and position must be given to both. With careful construction, interval response scales provide data that the majority of statisticians agree meet the assumptions for parametric analysis. 10-cm visual analogue scales provide similar sensitivity to within-patient change over time to the seven-point scale but they tend not to be quite so reliable and data extraction is often more difficult. The numbers, the verbal descriptors and the relationship between them should not be changed.

TIME SPECIFICATION

Most questionnaires have been validated using either a 1- or 2-week recall period, beyond which accuracy tends to deteriorate. Accuracy may also deteriorate with shorter recall periods, usually because patients’ experiences at weekends differ from those during the week (unless one is using a daily diary or a questionnaire for a rapidly changing condition [6]). The concept of time develops late in children and few under the age of 6 yrs can conceptualise “during the last week” [7]; hence the reason there are very few validated questionnaires completed by young children. The time specification should not be changed beyond the range for which it has been validated.

FORMATTING

Studies have provided developers with the knowledge of how to optimise formatting to enable patients to read instructions, questions and response options carefully, completely and accurately. In some languages, even the positioning of line breaks within a question will alter its meaning. Although a validated format may sometimes not look artistically elegant or match other formats in a case record form, it does give accurate results and should not be changed. Even changing from “circling the number” (patients’ preference) to “ticking the box” (case record makers’ preference) can affect how patients respond.

CULTURAL ADAPTATION

When questionnaires are adapted for other languages, the process is much more complicated than doing a simple translation. It has to be done by experts following recognised guidelines to ensure that the instrument is appropriately adapted for the local culture and climate, meets the original specifications for the instrument and that the measurement properties remain the same as those of the original [8]. I only allow cultural adaptations to be performed by a single institution. Not only do their translations follow recommended guidelines, they meet the requirements of regulatory agencies.

Dept of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada.

CORRESPONDENCE: E.F. Juniper, 20 Marcuse Fields, Bosham, West Sussex, PO18 8NA, UK.
E-mail: juniper@qoltech.co.uk

By using only one organisation, I can ensure international harmonisation (e.g. 14 countries use the Spanish Asthma Quality of Life Questionnaire, each one is slightly different but all 14 Spanish versions have been harmonised).

ELECTRONIC DATA CAPTURE

A more recent concern is the adaptation of paper questionnaires for data collection by electronic media (e.g. PCs, PDAs, phone, internet *etc.*). Two recent studies have shown significant bias and inadequate concordance between the original paper and the electronic versions (the first used a PDA, the second used two interactive voice response systems) [9, 10]. These failures were developed with as much care and testing as has been used for other devices that give valid data [11, 12]. The reason for some electronic versions performing poorly remains obscure and so there needs to be further research on this important issue.

CORRECT POPULATION

A questionnaire should be used only in the population for which it has been developed and validated. Inappropriate use usually occurs because the illness or age group does not have its own questionnaire (e.g. a rhinitis questionnaire should not be used in sinusitis; school-age children's questionnaires should not be used in infants; a disease-specific questionnaire validated in mild-to-moderate illness should not be used in severe patients).

CORRECT PURPOSE

Questionnaires are designed for different purposes, usually discriminative, evaluative, diagnostic or predictive [13]. The first are used to discriminate between patients of different levels of impairment (e.g. present/absent, mild/moderate/severe). The second are designed to measure within-patient change over time (responsiveness); these are the ones most commonly used in both clinical practice and research. Diagnostic questionnaires are, as their name suggests, specifically for identifying the presence or absence of an illness. Predictive instruments are used to identify likely outcome. Although instruments designed to have good evaluative properties often have good discriminative ones, the opposite is rarely true. Similarly, an instrument that has been developed for group analysis should be used with care in individual patients. For instance, the three shortened versions of the ACQ [5] were validated for large research studies and no longer have complete content validity for measuring asthma control in individual patients. Diagnostic questionnaires [14] go through a very different validation process. They should not be used for measuring outcomes and neither should outcome questionnaires (discriminative, evaluative, predictive) be used for diagnostic purposes.

CORRECT RESPONDENT AND LOCATION

Questionnaires should be completed by the same population as completed them in the validation studies. For instance, there is strong evidence that parents often have a poor perception of both their child's (aged >6 yrs) asthma status and asthma-specific quality of life [15] and so our children's questionnaires require the children themselves to be the respondents. Wherever possible, questionnaires should be completed in a similar location to the one in which they were tested (usually

the clinic). Home completions may risk biased responses, family interference, someone else completing the questionnaire, the questionnaire not being completed at the right time *etc.* It is wise to check alternative location of completion for validity [16].

CONCLUSION

As the authors of the PAX study have admitted, their unauthorized, modified version contained an additional question regarding the presence of a cold and a reformatted response scale for each question [17]; there were also some changes to the instructions and questions. It has been a long, hard struggle to get clinicians, academics, regulatory agencies and commercial companies to accept that subjective health status can be measured accurately and with precision. However, the struggle has been worth it because we now have a number of carefully developed and validated questionnaires and diaries that are proving invaluable in the assessment and management of patients' health and which are used as primary outcomes in research studies. If rogue versions get into circulation, confidence in the usefulness of these questionnaires will evaporate very quickly. Therefore, it is beholden to each one of us to ensure that we use only authorised versions in our clinical practice and research. Validated questionnaires and diaries are copyrighted to ensure that they are not altered, translated or adapted for another medium without permission. International copyright laws and intellectual property rights must be upheld for the well-being of patients.

STATEMENT OF INTEREST

A statement of interest for E.F. Juniper can be found at www.erj.ersjournals.com/misc/statements.dtl

REFERENCES

- 1 Juniper EF, O'Byrne PM, Guyatt GH, *et al.* Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999; 14: 902–907.
- 2 Walter MJ, Castro M, Kunselman SJ, *et al.* Predicting worsening asthma control following the common cold. *Eur Respir J* 2008; 32: 1548–1554.
- 3 Juniper EF, Guyatt GH, Streiner DL, *et al.* Clinical impact *versus* factor analysis for quality of life questionnaire construction. *J Clinical Epidemiology* 1997; 50: 233–238.
- 4 Stewart AL, Hays R, Ware JE. The MOS short-form general health survey. Reliability and validity in a patient population. *Med Care* 1988; 26: 724–732.
- 5 Juniper EF, Svensson K, Mörk AC, *et al.* Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. *Respir Med* 2005; 99: 553–558.
- 6 Juniper EF, Svensson K, Mörk A-C, *et al.* Measuring health-related quality of life in adults during an acute asthma exacerbation. *Chest* 2004; 125: 93–97.
- 7 Juniper EF, Guyatt GH, Feeny DH, *et al.* Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. *Eur Respir J* 1997; 10: 2285–2294.
- 8 Acquadro C, Conway K, Giroudet C, *et al.*, Linguistic Validation Manual for Patient-reported Outcomes (PRO) Instruments. Lyon, MAPI Research Institute, 2004.
- 9 Juniper EF, Schatz M, Juniper BA, *et al.* Evaluation of two interactive voice response telephone versions of health-related quality of life questionnaires. *J Allergy Clin Immunol* 2008; 122: 654–655.

- 10 Juniper EF, Langlands JM, Juniper BA. Patients may respond differently to paper and electronic versions of the same questionnaires. *Respir Med* 2009; 103: 932–934.
- 11 Caro JJ, Caro I, Caro J, *et al.* Does Electronic implementation of questionnaires used in asthma alter responses compared to paper implementation. *Qual Life Res* 2001; 10: 683–691.
- 12 Juniper EF, Riis B, Juniper BA. Development and validation of an electronic version of the Rhinoconjunctivitis Quality of Life Questionnaire. *Allergy* 2007; 62: 1091–1093.
- 13 Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 1992; 45: 1341–1345.
- 14 Price DB, Tinkelman DG, Halbert RJ, *et al.* Symptom-based questionnaire for identifying chronic obstructive pulmonary disease in smokers in smokers. *Respiration* 2006; 73: 277–278.
- 15 Guyatt GH, Juniper EF, Feeny DH, *et al.* Children and adult perceptions of childhood asthma. *Pediatrics* 1997; 99: 165–168.
- 16 Pinnock H, Juniper EF, Aziz S. Comparison of postal *versus* supervised responses to the Mini Asthma Quality of Life and Asthma Control Questionnaires. *J Clin Epidemiol* 2005; 58: 809–814.
- 17 Walter MJ, Castro M, Kunselman SJ, *et al.* Author correction: “Predicting worsening asthma control following the common cold.”. *Eur Respir J* 2009; 34: 1212.