Adapting Quality of Life Instruments

Angelo Swaine-Verdier, MA,[1] Lynda C. Doward, BSc,[2] Peter Hagell, PhD,[3] Hanne Thorsen, MD, PhD,[4] Stephen P. McKenna, PhD[2]

[1]Freelance translator, Chevalet, Mazerat-Aurouze, France; [2]Galen Research, Manchester, UK; [3]Department of Nursing, Lund University, Lund, Sweden; [4]Department of General Practice, University of Copenhagen, Copenhagen, Denmark

ABSTRACT

Due to the international nature of many clinical studies and trials it is often necessary to produce several language versions of specific measures. While it is generally acknowledged that it is necessary to produce versions that are conceptually equivalent, the best method of achieving this is more controversial. It is commonly stated that there is a gold-standard method, which involves forward and backward translation. However, no evidence has been presented to support this view. This paper argues that the "gold-standard" method is difficult to support and describes an alternative method involving dual translation panels that has been used in the production of all adaptations of needs-based quality of life instruments.

Keywords: construct validity, cultural equivalence, reproductivity, responsiveness, unidimensionality.

Introduction

Most health-related quality of life (HRQL) and quality of life (QoL) measures were first developed in UK or US English but clinical trials and comparative studies require instruments in several languages. This leads to consideration of the methods required to produce translations of high quality.

Early attempts to adapt measures sometimes failed possibly because the task is deceivingly simple. Researchers may have translated a questionnaire on a corner of their desks or turned to a professional translator. In the first case, linguistic ability may have fallen short of that required and in the other, a technical awareness of what is involved was often lacking. When translation work was given to "professionals" the need was felt to assess the quality of the target version by some sort of "scientific" method. This is probably how the back-translation method arose and it has become a very widely recommended methodology [1–5]. Despite this it has no clear scientific basis.

Back-Translation

The basic method consists of obtaining one or several forward translations by independent translators. Other translators then back-translate an agreed version into the source language. The two versions in the source language are then compared and "discrepancies" highlighted. Where several translations and back-translations are performed personal experience indicates that this comparison process, often performed by a consensus group, becomes unmanageable due to the sheer volume of text generated.

Such back-translation raises the hairs on the necks of translators and not merely because it casts unwelcome doubts on their abilities. If the translation is good, the back-translation may look nothing like the source questionnaire. Consequently, little information of any value will be obtained from a back-translation, while a lot of misleading impressions may result. While this is intuitive to any competent translator, it can also be demonstrated:

Situation 1: The Forward Translation Is Good

If the translation is to take into account all the requirements it may be very different in form and in apparent content from the source item. Hence, if there is then a good back-translation it is quite likely to differ considerably from the original. Thus comparing the two source language versions may not inform greatly on the quality of the translation, unless the comparison is made by someone who has a good enough command of the target language to see that the match is in fact there. Since those testing the quality of the translation need to do so precisely

Address correspondence to: Lynda Doward, Galen Research, Enterprise House, Manchester Science Park, Lloyd Street North, Manchester M15 6SE, UK. E-mail: ldoward@galen-research.com

because they do not have expertise in the target language, the exercise is of questionable value.

Situation 2: The Forward Translation Is Good, but the Back-Translation Is Poor

There is no reason to suppose that back-translators are better at their job than forward-translators, yet this assumption is made when they are used to test the forward version. In this case the back-translation may suggest that the forward version is poor. At best, such a situation is unproductive of useful information.

Situation 3: The Forward Translation Is Poor

The translation may be too literal or stay too close to formal aspects of the original in sentence structure or in choice of words. What can happen here—the most damaging situation—is that the back-translator, doing his or her best to produce something readable in the source language, finds the exact wording of the original item. This would not happen by chance. Poor translations are easy to identify because the forms and words of the source language pervade the translation. Consciously or unconsciously the translator will return to the source wording. This means that the back-translation, supposed to assess the quality of the translation, will mistakenly indicate that it is good. A number of researchers have highlighted problems with this approach [6,7].

Producing Quality

The answer must be to produce quality *in* the translation, rather than checking it a posteriori [8–10]. Producing quality requires checking and rechecking throughout the process, as well as after it, to see if the instrument functions as required with "real" people.

The following recommendations are made in order to ensure high quality adaptations:

- Recruiting translators with varied profiles to work as a team in a group meeting. A group of five to seven enables fruitful discussion.
- Informing the group of the model underlying the questionnaire, how it was developed, its design and content, and issues such as the target audience.
- Informing them of the translation requirements (in particular, conceptual equivalence, accessibility, and acceptability of wording).
- Having them work under an experienced coordinator whose task is to check that none of the parameters are neglected.
- Once the translated version of the instrument has been agreed, having it assessed by a lay panel working as a focus group in the target language only. The group should have access to only the translated version (not the source version) of the questionnaire as their assessment should not be affected by what they think the translated items *should* mean but rather what they do mean. The coordinator involved in the first panel should work with this panel to ensure that the original meaning of items and questionnaire structure are maintained.
- The whole procedure should be reported in detail, in particular explaining translation choices and changes made following lay panel testing. This not only informs the instrument authors but also constitutes a thorough final review.

It is important to remember that translation is only the start of the adaptation process. The agreed translation should then be tested by means of one-to-one interviews with several (typically 15–20) relevant representatives of the target population. This is done to ensure linguistic, face, and content validity. Finally, the psychometric properties of the adapted questionnaire should be established with new patient samples. In addition to establishing the instrument's psychometric properties in its new target population, this process should also assess the presence of differential item functioning (DIF) between the target and source versions [11]. Demonstrated absence of DIF indicates measurement equivalence and is a prerequisite for valid combination and comparison of data from the two language versions [12].

Translator Profiles

Since the questionnaires are to be read and used by ordinary people, it is preferable for translators to be as "ordinary" as possible. However, translators tend not to be ordinary but well educated. This can lead to their adopting an over-academic style, leading to translations that may be inappropriate for subsequent respondents. This problem can be countered by involving translators who are involved in activities that bring them into contact with a wide variety of people and language abilities. The most obvious example is teachers, who are well aware that wordings are crucial to the transmission of information.

It is helpful to have at least one native speaker of the source language involved in the translation panel to inform on the precise meaning of items in the source instrument. Frequently this will be one of the instrument developers.

The inclusion of a medical practitioner is unnecessary, as the source questionnaire should have overcome any technical issues. Similarly, there should be no need to include a researcher involved in questionnaire design in the translation process. A coordinator with access to background information about the questionnaire and sufficient experience in the field will be able to deal with any such issues and complex questionnaire formats should be discouraged. The coordinator needs command of both languages to be in a position to monitor the whole proceedings, make final decisions, and produce a meaningful report.

Lay Panel Function

This step is an integral part of a quality translation procedure. Five to seven people should be included to ensure that the group is varied enough, while allowing adequate discussion. Ideally, participants should meet the following criteria:

- There should be a balance between the genders and age groups unless the target population for the instrument is biased by gender or age.
- It is preferable for the participants to be of average educational attainment in relation to the population of interest, with some participants of a fairly low educational status included in order to test the questionnaire adequately for comprehensibility.
- They should represent varied occupational and social backgrounds.
- Individuals with the disease covered by the questionnaire should be omitted. This is because they tend to be too concerned about the suitability of the questionnaire content (which cannot be changed) and too little about the appropriateness of the wording. The next stage of the adaptation process involves testing with relevant patients.

The focus group procedure consists of a brief explanation of the reasons for the meeting and what is expected of the group members. After this, participants discuss the items as a group, rewording them where it is considered necessary. This process allows a check on whether the appropriate concepts have been captured and if individual items are comprehensible and acceptable in content and wording. Other outcomes are also possible:

- Wordings that the first panel were searching for but were unable to find are frequently identified.
- Other wordings considered questionable can be shown to be acceptable and well understood.
- Unforeseen connotations may be identified. Translators may not be aware that members of the lay panel can unexpectedly and wrongly associate words with certain ideas. Such associations are a reality due to cultural differences in conceptual meaning and should be taken into account.

Reporting

A detailed report on the translation procedure is essential. It should explain why changes in form or content were made, why some items were difficult to translate, where cultural issues have to be addressed, why (for example) a word such as "embarrassed" cannot generally be translated as "embarrassé" in French and other explanations of why "literal" translations, that may seem obvious to anyone with an approximate knowledge of the language, are not suitable. Rough translations can be used to illustrate these points but such translations will not necessarily be in a form that would be suitable as source language items—this is not their purpose. Such explanations are also a basic courtesy to the instrument's authors.

The report is also an important tool for the users of the translated instrument in the subsequent validation phases. It identifies issues that need to be tested further during patient interviews and may help to explain why items fail to function as expected.

Conclusion

These comments are intended to contribute to the debate on instrument translation and adaptation. A review of the literature would suggest that there is little alternative to using back-translation but the "two-panel" approach described above has been widely and successfully employed in the adaptation of needs-based measures of QoL [13–18].

Translating is never a straightforward process, though those who do not practice it may view it as a mathematical task of finding "equivalents" and stringing them together. A given language is a specific way of putting life into words; and words change with their context. They generate representations that may be universal or culture-specific and some representations may be unintended or even upsetting. Therefore, it is often more appropriate to consider the process as one of adaptation rather than translation. The procedures described above are not limited to cross-lingual instrument adaptations but are also necessary for valid use of UK instruments in American or Canadian English speaking populations and vice versa [13,15,17]. Words are particularly emotionally loaded when they are used in an area as essential and personal as QoL or even HRQL. Using an instrument with such a population requires a particularly important dimension of respect and ethical responsibility. That is why, in addition to the need to produce an efficient and accurate tool, the development of a translated version should also ensure acceptability on ethical grounds.

References

1 Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–32.
2 Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. Qual Life Res 1993;2:451–9.
3 Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. J Clin Epidemiol 1998;51:913–23.
4 Lent L, Hahn E, Eremenco S, et al. Using cross-cultural input to adapt the Functional Assessment of Chronic Illness Therapy (FACIT) scales. Acta Oncol 1999;38:695–702.
5 Norholm V, Bech P. The WHO Quality of Life (WHOQOL) Questionnaire: Danish validation study. Nord J Psychiatry 2001;55:229–35.
6 Fukuhara S, Bito S, Green J, et al. Translation, adaptation, and validation of the SF-36 health survey for use in Japan. J Clin Epidemiol 1998;51:1037–44.
7 Hagell P, McKenna SP. International use of health status questionnaires in Parkinson's disease: translation is not enough. Parkinsonism Relat Disord 2003;10:89–92.
8 Hunt SM, Alonso J, Bucquet D, et al. Cross-cultural adaptation of health measures. Health Policy 1991;19:33–44.
9 Leplège A, Verdier A. The adaptation of health status measures: methodological aspects of the translation procedure. In: Shumaker S, Berzon R, eds., International Use and Performance of Health-Related Quality of Life Instruments. Oxford: Oxford Rapid Communication, 1995.
10 McKenna SP, Whalley D. Coping with language barriers. Good Clin Pract J 1997;4:14–17.
11 Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development of quality of life measures. Value Health 2004;7(Suppl 1):S22–S26.
12 Cella DF, Lloyd SR, Wright BD. Cross-cultural instrument equating: current research and future directions. In: Spilker B, ed., Quality of Life and Pharmacoeconomics in Clinical Trials 2nd ed. Philadelphia: Lippincott-Raven Publishers, 1996.
13 Doward LC, McKenna SP, Kohlmann T, et al. The international development of the RGHQoL: a quality of life measure for recurrent genital herpes. Qual Life Res 1998;7:143–53.
14 McKenna SP, Doward LC, Alonso J, et al. The QoL-AGHDA: an instrument for the assessment of quality of life in adults with growth hormone deficiency. Qual Life Res 1999;8:373–83.
15 McKenna SP, Doward LC, Kohlmann T, et al. International development of the Quality of Life in Depression Scale (QLDS). J Affect Disord 2001;63:189–99.
16 Doward LC, Spoorenberg A, Cook SA, et al. The Development of the ASQoL: a quality of life instrument specific to Ankylosing Spondylitis. Ann Rheum Dis 2003;62:20–6.
17 Whalley D, McKenna SP, Dewar AL, et al. A new instrument for assessing quality of life in Atopic Dermatitis: international development of the Quality of Life Index for Atopic Dermatitis (QoLIAD). Br J Dermatol 2004;150:274–83.
18 Thorsen H, Hansen TM, McKenna SP, et al. Adaptation into Danish of the Stanford Health Assessment Questionnaire (HAQ) and the Rheumatoid Arthritis Quality of Life Scale (RAQoL). Scand J Rheumatol 2001;30:103–9.