Quality of Life Research (2006) 15: 1257–1270
DOI 10.1007/s11136-006-0062-4

Issues in conducting cross-cultural research: implementation of an agreed international procotol designed by the WHOQOL Group for the conduct of focus groups eliciting the quality of life of older adults

Graeme Hawthorne[1], Natasha Davidson[1], Kathryn Quinn[2], Farah McCrate[3], Ines Winkler[4], Ramona Lucas[5], Reinhold Kilian[6] & Anita Molzahn[7]
[1]*Department of Psychiatry, The University of Melbourne, Melbourne, Australia ;* [2]*Department of Psychiatry, Edinburgh University, UK;* [3]*Department of Psychology, University of Bath, UK;* [4]*Department of Psychiatry, University of Leipzig, Leipzig, Germany;* [5]*Institut Catala de I'Envelliment, Universitat Autonoma de Barcelona, Barcelona, Spain;* [6]*Department of Psychiatry II, University of Ulm, Günzburg, Germany;* [7]*School of Nursing, University of Victoria, Victoria, Canada*

Accepted in revised form 13 March 2006

Abstract

Multi-centre and cross-cultural research require the use of common protocols if the results are to be either pooled or compared. All too often adherence to protocols is not discussed in reports and where it is reported poor adherence is frequently noted. This paper discusses the use of international guidelines developed by WHOQOL Field Centres to conduct and report focus groups aimed at eliciting key concepts of quality of life among older adults. This was the first step in the development of the WHOQOL-OLD instrument. Although there was overall adherence to the agreed guidelines, there were some differences in the level of reporting, even after participating Field Centres had the opportunity to explain their reports. The reasons for these discrepancies are reported. It is concluded that because of local situations, it is difficult to achieve identical implementation of multi-centre cross-cultural protocols and that the highest standards of auditing are required if findings are to be compared. Suggestions for how such protocols can be improved are given.

Key words: Cross-cultural research, Instrument development, Multi-centre trials, Quality of life, Research protocols, WHOQOL, WHOQOL-OLD

Introduction

As the world community becomes increasingly integrated, diseases and the results of social changes, including political unrest, spread more quickly than ever before [1]. Health interventions follow the same pathway. As a result, more multi-centre cross-cultural studies are being conducted. There is a need for systematic monitoring and reporting of the experiences of researchers and outcomes related to the research process [2]. An issue of growing importance is cultural difference and the extent to which this affects health interventions and reported outcomes [3]. If cross-cultural multi-centre interventions and studies are to possess generalizable validity, then researchers need to ensure common research procedures are followed and that the measures used have the same meaning in different cultures [4, 5].

One obvious area is in relation to the measurement of quality of life (QoL). As QoL measures become more widely used, cross-cultural validity becomes a challenge for instrument developers [4, 6, 7]. Cross-culturally valid measures are important as they can provide information enabling cross-

cultural comparisons at the descriptive level (e.g. comparing the health or QoL status of different populations or cultural groups) as well as at the inferential level (e.g. comparing the effect of a particular treatment on different groups). They also enable QOL comparisons across different cultural or social groups where, in a range of settings, there may be different forms of health care [2].

Generally, however, where QoL instruments have been developed they reflect the values and concerns of clinicians, patients and the general public of the country of origin [8]. For example, regarding health-related multi-attribute utility measures, instruments have been developed in the UK, US, Finland, Canada, the European Union and Australia. Other than the EQ5D, which was developed in the European Union but weighted with British values, none of these instruments has been validated for use in cross-cultural settings; this validity is simply assumed [9]. This assumption is not necessarily warranted as can be seen in the case of the SF-36, the world's most ubiquitous health status measure: there are differences in wording and scoring between different country versions [10], even where the language is common [11].

To ensure cross-cultural validity, during instrument construction and validation multinational collaboration is required at three levels. The conceptual latent model underpinning an instrument's manifest model must reflect different cultural nuances [2, 5]; the manifest model – an instrument's descriptive system – must measure, in a representative sense, the universe of interest as defined cross-culturally in order to take account of emic and etic effects [12]; and the observed model of QoL elicited by the instrument must be shown to have validity in different cultural settings [13].

A key reason why validation at these three levels is rarely performed is that there are recognised challenges associated with conducting multinational and multi-centre research [14]. Problems arising with international adherence to protocols have been reported [15, 16]. There is little published information in relation to the processes involved with this collaboration and the problems of embarking on large scale, multi-centre cross-cultural research. One group with wide international representation that is tackling this difficult area is the World Health Organization's QoL group [WHOQOL Group, 14, 17]. To date the WHOQOL Group has used a common international protocol to develop two generic QoL profile measures, the WHOQOL-100 and the WHOQOL-BREF [18–20]. Although these instruments are now widely used, the WHOQOL Group has not published on issues concerning the challenges inherent in conducting the multi-country research leading to the development of these instruments.

The WHOQOL Group is currently developing a third generic instrument, the WHOQOL-OLD which is explicitly designed to measure, cross-culturally, the QoL of older adults. This project involves 23 Field Centres from around the world, spanning a wide range of cultural types (e.g. from Western European countries such as the United Kingdom and Spain, to Asian countries such as Japan and China). Within each participating country, the study is managed by a local Field Centre. A common international set of guidelines for instrument development was prepared centrally at the coordinating university, Edinburgh University. The guidelines involve focus groups (to elicit the universe of interest), development and piloting of an international item bank, participating in instrument construction procedures and conducting a validation study. For each of these study phases, detailed guidelines were prepared and each participating WHOQOL Field Centre was expected to implement these so that the data were common across all Centres and could be pooled for instrument development or validation work. Clearly the first stage – the conduct of focus groups – was critically important to the whole process.

Focus groups represent an important qualitative tool for discovery and exploration when little is known about a particular topic. Successful focus groups enable information to be gathered in a non-directive manner by allowing discussion to flow naturally and they enable elicitation of a wide range of views on the subject of interest [21]. Because of these two important attributes, focus groups are widely used in research to provide exploratory and detailed information about topics of interest, which can then be used to inform the further development of research studies and instruments. One area where focus groups have been successfully used is in QoL research; studies have included breast cancer survivors, patients with osteoporosis, dialysis patients, and those with

migraine and diabetes [22–26]. Focus groups have also been used to develop an understanding of the dimensions of QoL where this information has then been used for guiding QoL instrument development [27].

This paper reports on the processes associated with the implementation of the WHOQOL-OLD focus group guidelines. The aim of the study was to examine how 19 of the participating Field Centres implemented these for the first phase of the WHOQOL-OLD development (the exclusion of the other 4 Field Centres was because they joined the study at a later phase). The particular concern was to document the issues involved in conducting cross-cultural research where common international guidelines were previously agreed by the research partners.

Methods

The WHOQOL-OLD Coordinating Field Centre produced draft guidelines based on the previous WHOQOL Group experiences in conducting international collaborative research for the development of the WHOQOL-100 and WHOQOL-BREF [18–20]. Following initial development, the guidelines were circulated to each Field Centre for comment. They were iteratively revised using a Delphi technique until there was agreement among the participating Centres. The intent was that the guidelines would facilitate consistent data collection and reporting of focus groups across all Centres, although it was stated that participating Centres could change the guidelines for running the focus groups to suit their particular circumstances. The guidelines also established a common framework for interpreting and assessing the data reported by each Centre. Once agreed, they were used in each Centre to plan and conduct focus groups for the purpose of eliciting the universe of QoL concerns of older adults, and for reporting the data back to the Coordinating Centre.

Reportable task list from focus group guidelines

The guidelines covered all the various tasks associated with running the focus groups and provided a suggested structure for the conduct of these. There were 13 reportable tasks, which for this study are reported under 7 general areas. These were used in this study as the criteria for assessing how well Centres implemented the guidelines. These were:

1. Timeframe for conducting the groups, the appropriate sites in which to conduct the groups, transport to and from the focus groups, and the timing of the groups.
2. Number of focus groups, recruiting of participants and group composition, including inclusion and exclusion criteria; i.e. the methodology for setting up the focus groups. The guidelines suggested that there should be 4–6 participants in each group of older adults, carers and health professionals. The older adult groups were to be organised by age ($<80/80+$ years) and health status (healthy/unhealthy). Groups were to be of mixed gender.
3. Preparation for conducting the groups, including translation of materials, pre-distribution of stimulus materials and information for participants (optional), and the materials and equipment needed during the groups.
4. Procedure for conducting the groups, including the role of the moderators, the structure of sessions, and advice on conducting focus groups with older adults. This advice included material on meeting with the participants, tips for getting started, how to best observe and record the groups and closing the groups. Information on how to handle difficult situations was also included.
5. Socio-demographic and health status information to be collected and reported. This referred to the age of participants, their general health status (healthy/unhealthy), medical conditions, education attainment, use of medications and current living circumstances.
6. Structure of the focus group discussion. This involved four parts: a general unstructured discussion on the dimensions of QoL that were important to older adults, to comment on and assess facets and items from the WHOQOL-100 instrument [19], to give feedback on additional facets and items that had been previously suggested by Field Centres during the Delphi exercise described above, and to gather ideas from participants for additional areas of QoL or items

that participants felt were not covered during discussion.
7. Translation of focus group findings, and the documentation required in the reports to be forwarded to the Coordinating Centre.

In general, the guidelines outlined a model which followed general protocols for successful focus group implementation [28, 29]. A semi structured approach was to be used to ensure that core concepts of QoL were covered across the groups conducted within each Field Centre, and that issues particular to each group could be explored. Each Centre agreed to conduct four focus groups with older adults, one with their carers and one group with health professionals working with older adults (i.e. a minimum of six focus groups).

Participants

Focus group reports from 19 Field Centres were analysed for this study. The Centres were Barcelona (Spain), Bath (England), BeerSheva (Israel), Budapest (Hungary), Copenhagen (Denmark), Edinburgh (Scotland), Guangzhou (China), Hong Kong (China), Leipzig (Germany), Melbourne (Australia), Montevideo (Uruguay), Oslo (Norway), Paris (France), Porto Alegre (Brazil), Prague (Czech Republic), Seattle (USA), Tokyo (Japan), Umea (Sweden) and Vilnius (Lithuania).

Method for establishing adherence to protocol

A matrix was compiled by listing from the guidelines each of the 13 criteria for conducting the focus groups. Each Centre's report was examined against the criteria. These data were then summarised to produce a description of adherence to the guidelines, using a 3-point scale for each criterion (2 = fully complied with the criterion, 1 = partly complied, 0 = did not comply). Hereafter this is referred to as ''report'' analysis. The results were then returned to each Centre for comment on the accuracy of the coding, clarification on the areas of their report that were unclear, and verification of the reasons why it did or did not adhere to the guidelines; this is referred to as ''verification''. The 13 categories of information were then collapsed into the 7 criteria reported in this paper. The amalgamation was based on identifying similar criteria, for example the three requirements that Centres report the demographics of participants, carers and clinicians were collapsed into reporting demographics.

These processes allowed examination of how the Centres adhered to the agreed focus group guidelines and to observe any systematic differences that may have emerged. Themes, issues or problems emerging in adherence to the guidelines across Centres were then extrapolated and described. Reasons for the deviation from the guidelines were described in the light of conducting cross-cultural and multi-centre studies.

The reports were collated using Excel [30], NVivo [31] was used to code the data within reports, and the data were analysed using SPSS [32] and InStat [33].

Results

There was great variation in the reports from the participating Centres. Some reports were long and detailed; the longest was over 100 pages. Others were very short and concise; the shortest was 2 pages. Essentially this meant that very different quantities and levels of data were available from the various Centres. Subject to these differences, basic data about the focus groups are presented in Table 1.

This shows that all Centres conducted the required 4 focus groups with older adults, except for Tokyo which conducted three. Most Centres conducted two carer focus groups, and one clinician group. Regarding the participants, 68% were living at home, 18% were living in a hostel or other supported accommodation and 14% were in hospital as either an inpatient or outpatient. For education level, the table shows most had achieved secondary education (46%), followed by university (28%) and primary level education (27%).

On the 0–2 coding scale described in the methods section, the mean across all major categories was 1.49 (sd = 0.21) for reports and 1.71 (sd = 0.18) for verification. Although most of the Centres included the required data in their reports, these differences suggested that Centres did not report all the information they collected. When explicitly asked about the individual criteria, they were able to provide more information.

Table 1. Details of focus groups and participants

	N. Focus groups conducted			Cases living arrangements			Cases education level		
	Cases (a)	Carers	Clinicians	Home	Hostel	Hospital (b)	Primary	High	University
Barcelona	4	2	1	17	0	0	14	3	0
Bath	5	1	1	27	4	0	0	17	9
Beer-Sheva	4	2	1	6	16	0	9	11	2
Budapest	6	1	1	21	10	5	13	8	10
Copenhagen	4	2	1	26	3	0	6	15	8
Edinburgh	4	2	1	10	0	10	0	17	2
Guangzhou	4	2	1	17	0	4	12	8	1
Hong Kong	4	2	1						
Leipzig	4	1	1	13	4	0	2	4	11
Melbourne	4	2	1	20	8	0	5	22	1
Montevideo	4	1	1	5	5	7			
Oslo	4	2	1	11	0	11	4	4	11
Paris	6								
Porto Alegre	4	1	1	0	11	4	11	3	4
Prague	4	1	1	22	1	5	0	16	9
Seattle	4	2	1	15	4	0	0	8	11
Tokyo	4	2	1	18	0	11	3	16	10
Umea	4	2	1	15	0	0	6	6	3
Vilnius	4	1	1	11	7	0	1	9	8
Total	76	29	17	243	66	26	85	158	92

a = Older adults.
b = Includes both outpatients and inpatients.

To explore which areas were less well reported, the ratings for each of the seven criteria were examined and the report and the verification scores compared. This showed that timeframe was more completely reported and the reporting of the groups was less well reported in both the report and verified scores.

Differences were also observed in relation to the preparation for the focus groups and reporting of demographic information. In both instances verification scores indicated better compliance with the guidelines than did the report scores: the medians were 2.0 and 1.5 for preparation and 2.0 and 1.7 for demographics respectively. Where there were differences between the report and verification scores, Centres were asked to comment on the reasons for these. Reports from the Centres regarding the differences follow.

Timeframe

Eighteen Centres adhered to the stated timeframe in the guidelines, which was that focus groups should last between 1.5 and 2 hours, with appropriate breaks for refreshment. Following verification, the timeframe was partially completed in one Centre. The reason for this was that:

> ...participants complained about the length of the questionnaire and moderators decided not to persist.

Consistent with this, several Centres reported that towards the end of the focus groups, participants were noticeably tired.

Number of focus groups and methodology for setting up focus groups

The guidelines suggested 4–6 participants of mixed gender in each group of older adults, carers and health professionals. The older adult groups were to be organised by age ($<80/80+$ years) and health status (healthy/unhealthy).

Centres were asked to clarify the focus groups composition since 12 of 19 Centres had partly followed the guidelines. Of these 12, six stated in their report that the guidelines had been adhered to. Seven Centres confirmed they had only partly adhered to the guidelines. The reasons for this

mainly centred around the difficulties in self determination and moderators' classification of the health status of older adults. For example:

> In relation to the perceived health status – this is something it has been raised several times – we ran two FG in Primary Care and two FG in Community Centres thinking that PC participants will report themselves as sick, and Community participants as healthy. But, the real finding is that people reporting some illness perceived themselves as healthy. So, we ended with mixed groups. I believe that this is very important and it is an issue we have raised several times in our meetings.

> (80 plus unhealthy) This is the group that was being referred when it was stated that group composition was more difficult to control. It took place in a sheltered housing scheme. The members were all women and they had various problems associated with age related decline. The oldest member of the group was 94 years old. Due to their various problems they were considered unhealthy and needed more or less the constant care/supervision of a sheltered home.

With regard to the criteria for mixed gender, all three groups (professional, carers and older adults) indicated an over representation of female participants (70% of all participants) because they were easier to access, their greater life longevity and willingness to talk about QoL issues. This was a common theme across many Centres:

> We tried to have gender mixed focus groups and it was fulfilled in four FGs with two exceptions. The first one was 80+ healthy group and group of professionals. There were no men visiting the daily club for seniors 80+. The second homogenous FG was the professional FG since only women worked as nurses or occupational workers in institutions for seniors that we contacted.

> Limitations in recruiting elderly males; more female longevity, as well as lack of willingness to discuss these issues openly on part of men.

> Females predominated despite attempts to organise groups of mixed gender; may reflect social trends towards increasing numbers of women in older age groups and attendance at day clinics.

> For almost all groups it was really difficult to find or to convince male participants. Some that were asked were not willing to participate because of different reasons (no time, didn't see the sense in it etc).

> In our report we stated that our older population has a higher percentage of women and also they show more disposition to talk about these topics.

Preparation for conducting groups

Preparation for conducting the groups included translation of materials, pre-distribution of stimulus materials if Centres wished to do so (although it was not necessary) (WHOQOL-100) and information to participants, and the equipment and materials needed during the groups. These were areas where the reports contained limited information. The inclusion/exclusion criteria were seldom reported; although at verification 18/19 Centres stated that they had adhered to the criteria none had stated fully meeting this criterion in their report. Five Centres referred to this criterion, but in 14 reports it was not mentioned at all. Two Centres commented that adherence to the criteria was assumed:

> The reason not to refer to this (inclusion and exclusion criteria) in our report was that we chose to report what was divergent from the guidelines.

> These exclusions were not noted in the report since the report was for internal use only, and adherence to these criteria was assumed.

Centres were asked to report if in advance of the focus groups they had distributed copies of the WHOQOL-100 questionnaire. In six reports there was information showing that this criterion was partially met; it was not mentioned in nine reports.

Following verification, 15 Centres confirmed they followed the guidelines. The reasons why this information was not included in the reports varied and reflected the diversity of circumstances in relation to conducting research to tight deadlines, participant expectations, organisational difficulties and the limitations of older adults:

> We did not distribute blank copies (of the WHOQOL-100) because of lack of time.

> We did not have the opportunity to do this since we did not know in advance who would be attending the session. However, we did give our participants some time to read through it during the actual session.

> Actually we had prepared for all groups the WHOQOL-100, since some of them were illiterate, therefore we did not distribute to those who could not read. Instead we used a poster format and told them about the details of the WHOQOL-100 during the session to facilitate discussion.

> Group 4 complained that it was too tiring to read the questionnaire beforehand.

> We have experienced that most persons couldn't held on attention properly, so later participants received sheets only after the first part of session during a break.

Regarding the preparation for the focus groups, Centres were asked to report their activities in relation to getting the moderators involved in setting up, translation of guidelines, providing transport and facilities as necessary. In the reports, one Centre reported it had complied with these criteria, in seven reports it was partly reported and it was not reported in the remainder. When asked, 17 Centres stated they had fulfilled these criteria, but no Centre gave reasons for the incomplete reporting.

Procedures used within the focus groups

The guidelines indicated that Centres should have provided a plain language statement and consent form at the start of each session and offered participants the chance to ask questions. This was reported in 13 reports. After verification, all 19 Centres confirmed they had fulfilled this criterion completely. One Centre acknowledged that this information was not recorded in the report:

> Our moderators had followed the suggested procedures in the sessions, maybe we did not indicate [this] clearly in the report.

Two Centres reported two very different reasons for not collecting written consent:

> We considered it was not necessary to report informed consents were collected, as if it was taken for granted.

> The participants were pretty suspicious about signing the consent forms because of the promised anonymity. So they only gave a verbal consent to us.

Socio-demographic information

Centres were asked about whether they collected and recorded demographic information on all participants.

First, regarding the older adults and their carers, this was not reported in 7 reports and it was partly reported in four reports. Two Centres gave reasons for this:

> We only reported age and health status of participants because we considered it was the main data needed in the reports.

> The reason for not addressing these issues was that we couldn't see the relevance of doing that. The important part was the content of the focus group discussion.

Following verification, 14/19 Centres confirmed that they collected these data, even if they had not reported it, and five Centres confirmed that they had collected some data. In general the reasons for omission were that this information was not considered important to the study.

For the health professionals, Centres were asked to report their demographic data including their professional qualifications. Thirteen Centres

reported at least some of these data in their reports. When asked, 16 Centres reported this information was collected but only partly reported because this was not a requirement of the guidelines. For example, one Centre stated that:

> The guideline indicated that we were required to collect the demographic data (age, sex, health status, etc.) of the elderly only.

It is noted that the guidelines stated that these data – for both the older adults and the health professionals - should be collected and reported. Thus it was not a requirement that these data were actually presented in the reports. Furthermore, the guidelines stated that these data could be collected at the end of the sessions. If participants became tired (as they did in several focus groups) or there was a time constraint, it may be that these data were perfunctorily collected.

Structure of focus group discussions

The guidelines suggested that focus groups should have four parts: a general unstructured discussion on the important dimensions of QoL, a review of the facets and items from the WHOQOL-100 instrument, a review of additional facets and items that had been suggested by Field Centres, and the gathering of ideas from participants on areas of QoL that were otherwise not covered during the discussion.

Centres were asked to comment on whether they used this suggested structure and format for conducting the focus groups with older adults and carers. The guidelines listed many requirements here, and it is possible this may explain why most Centres only partially fulfilled this requirement, even following verification. Sixteen Centres reported that they had partially adhered to the suggested format for focus group discussions. Of these, following verification five stated they had followed the guidelines completely, eight maintained that they were only able to partially adhere to these suggested criteria. The reasons for this were related to time constraints, the decreasing concentration level of participants and misinterpretation of what was requested:

> Because of the time constriction, we divided the WHOQOL-100 in two questionnaires with 50 questions each: pair (2,4,...100) and impair questions (1,3,5, ...99), and give one of each to each of the participants. In that way we assured that all questions were reviewed in the group.

> Not all facets were discussed in each group as it was anticipated there would not be sufficient time to do this. To accommodate for this, the facets were randomly divided in half and one half was covered in each group. After the first four groups, it became apparent which facets were more straightforward and which might be more problematic. Problematic facets or those that had not generated much discussion were highlighted for discussion in the later groups.

> The suggested questions were used, but not all questions were used in all groups. We did not guide the discussion into specific parts of the WHOQOL because the discussion evolved in a natural way into certain areas of interest.

> The structure was not followed fully with the group of 80+ unhealthy people. Because their concentration was limited, free form discussion of QoL (using the question) was predominant. They were asked about the QoL domains but the WHOQOL-100 was not discussed.

> There are no results to Part 1, as this was a free-form (warm-up) exercise designed to set the scene and to encourage participants to think about the issues most important to their QoL. In my recollection, we weren't required to report these findings.

In relation to specific items to include in the focus groups, the reporting of new items and their ranking was less likely as these seemed to be subsumed into the existing free form QoL and WHOQOL-100 discussion:

> Regarding the additional items... It did seem somewhat troublesome for the participants to make ranking of such a long list. They preferred to make oral statements.

> With regard to ranking suggested items it was impossible to rank 'intensity of opinion' as design of focus group discussion limited this.
>
> We did rank items in one group because it was appropriate to do so however in most groups suggestions were part of the general discussion and there was no time to rank these at the end.

Respondents were also asked to comment on whether they encountered problems with participants when conducting focus group discussions. The comments were to be in relation to sensitive issues, making participants feel valued and how valuable their opinions were to the study, reassurance of confidentiality, and maintaining open communication when there might be dominant members of the group. Nine Centres did not report these matters. At verification, however, although most Centres stated that although they were aware of problems, very few problems were identified:

> Re problems encountered in the groups. I recall noting on the form for one of the groups that we had difficultly discussing the more 'negative' issues – they preferred to talk about the good stuff. However, this wasn't a major issue. I don't recall any other problems encountered.
>
> Participants appreciated very much that finally somebody was interested in the elderly – we had no problems with individuals unwilling to speak, to participate.
>
> No problems, in fact there was great and plentiful dialogue.

For the health professional focus group, Centres were asked to use a different format to that used with the older adults and carers. These focus groups were to describe the use of the construct QoL in everyday work, the impact of QoL information on treatment choice, the perceived influence of treatment on QoL, the assessment of changes in patient QoL, and to review aspects of questionnaire structure and design. In two reports there was no record of focus groups conducted with professionals. However, following verification, both Centres affirmed they had conducted these and used the suggested format. In four Centres, these focus groups were conducted but there was no record of them covering the items. All four of these Centres stated they had partly met the guideline requirements. Seven Centres did not use all of the suggested guidelines for the following reasons:

> ...the QoL instrument is not common at all to be used in the clinical session as it said in the report.
>
> Some (items) were not addressed at all. E.g. questions about treatment options. Mostly the health professionals worked in aged care in community health, and were not medicos – therefore treatment was not highly relevant.
>
> We did run a focus group with health professionals but did not follow guidelines strictly as discussion flowed more spontaneously. We focused on the discussion of the questionnaire and additional items because of lack of time.
>
> The questions that were asked of the health professionals were taken from the guidelines, however, not all questions on the guidelines were utilised. This is because it was thought that it would take too long to get through all the questions as well as covering the designated facets of the WHOQOL-100 in the time allotted for a focus group. Therefore, the research officer and supervisor chose the questions they thought to be most useful and relevant to include.

Three Centres interviewed health professionals instead of conducting focus groups:

> We ran individual interviews with professionals due to difficulty arranging groups. We used a structured interview schedule covering all these questions and collated responses to them.

Translation and report of focus group findings

Centres had agreed that as part of the protocol they would document their focus groups. The information included the location and venue of

focus groups, group type (older adults/carers / professionals), duration, day/date/time, moderators involved, general atmosphere, nature of discussion flow, problems encountered/identified with focus group sessions and reasons participants left the focus group. Six Centres did not report these issues, and eleven partly reported them. Following verification, all 19 Centres reported using the guidelines template. One possible reason for this discrepancy is that the instructions were not fully clear: the guidelines asked Centres to return this information to the Coordinating Centre, but not necessarily include it in the report, thus:

> A note at the bottom of the form stated: ''Note for moderators: Please return this form, together with all other documentation, to the coordinating centre''. I did this at the completion of the groups. We were not requested to report this information in our focus group reports.

More importantly, Centres were asked to explain the process of documentation and translation of focus group discussions. This involved focus groups being recorded and then transcribed from audio tape, expressions used by participants to be reported and highlighting of relevant themes emerging from the focus group transcripts. In general, Centres transcribed the focus groups, but reported on general themes. Very few Centres quoted directly from the participants and even fewer highlighted emerging themes. This was a disappointing finding as the guidelines had stated:

> To facilitate item development, common items/phrases or themes should be highlighted together with any relevant comments.

The reports showed that six Centres had completely fulfilled this requirement, and a further four partly fulfilled it. After verification, eight Centres stated they had fulfilled the criteria. Reasons for not reporting this criterion are listed below:

> All sessions were recorded and transcribed except the staff interviews which were only summarised due to objection to recording.
>
> All focus groups were tape-recorded but the transcriptions were not verbatim.

Eleven Centres partly documented and translated focus groups transcripts, although six had not stated this in their report. Just one reason was given for this omission:

> This (transcription) was not possible because of lack of time as we entered the project later than other centres.

Discussion

In multi-centre and cross-cultural research challenges often emerge, including differences in language and communication, culturally sensitive interpretation of tasks, differential access to participants, availability of technology, resources and access to research funding. Yet, when multi-centre and cross-cultural research studies are reported, it is assumed that standardised research protocols are implemented. Differences in data collection and reporting raise questions regarding the comparability of the data and the validity of the findings. For example, if participants in a multi-centre trial are recruited differentially the study findings are unlikely to be comparable across the centres. The only way of knowing whether these types of difficulties have occurred or not is for research protocols to be followed and their implementation documented. Yet difficulty with poor adherence to protocol is a general problem in conducting multi-centre studies [16].

In this international multi-centre study, focus groups were used to elicit information about key areas of QoL relevant to older adults. The 19 participating Centres agreed to follow a set of international guidelines and to report on implementation in the local setting. In this paper, we compared each Centre's report with the guidelines, and where there were discrepancies we returned our findings to the Centre and asked if it could explain the reason for this.

The findings showed that the guidelines were generally followed. Where there were differences, many Centres reported that they had assumed that it was enough to complete the guidelines and not

report this, often because a criteria was deemed unimportant. This explanation, however, was not entirely satisfactory in light of the findings showing there were significant differences in how Centres adhered to and reported on the criteria. In particular, the timeframe criterion was well reported, whereas criteria relating to the structure of the focus group discussions and the translation and reporting of focus group findings were less well documented. Four Centres explained this by referring to objections to recording the focus group, that the recording was not transcribed and that there was insufficient time allowed for this.

Similar challenges included problems with group composition, especially with older adults (determining health status and age requirements), providing the appropriate gender mix, and problems associated with older adults' level of concentration during focus groups. This was reported both in the reports and during verification. Similarly there was some variation in the conduct of the health professional focus groups, including that some Centres conducted individual interviews rather than a focus group. Other Centres noted that some elements of the guidelines were not relevant to all professionals as these aspects of care were areas they were not involved with.

Many reasons could be postulated as to why the differences reported above occurred. Issues in translation and interpretation of the guidelines must be considered; it is possible translation techniques for both the guidelines and focus group findings may have varied and led to differing interpretations. This, however, must be balanced by the fact that all the lead researchers were able to communicate in English and that they were funded for translation and transcription. Another reason may be related to the use of the WHOQOL-100 in the focus groups; the instrument length caused some difficulties including running out of time, loss of concentration and the departure of some participants. Additionally, there were logistical problems, including organisational deficits (e.g. unavailability of staff).

It is likely another source of variation was within the guidelines themselves. Although all Centres had agreed to the guidelines through the Delphi technique, the guidelines provided a *suggested* structure for conducting and reporting the focus groups. This is particularly important because it implies that there was no required or mandatory standard of reporting. For several reportable tasks, the guidelines asked Centres to return this information to the Coordinating Centre, but not necessarily to include this in the report. As the results show, interpretation of what was required varied considerably. A third source of variation may be the very collaborative nature of the WHOQOL group whereby individual teams have high levels of local automony. It is possible that greater consistency could be obtained in studies where the central research team exercised a high level of supervision.

These observations are consistent with the differences in the length and detail of the reports. The average report length across the 19 participating Centres was 31 pages (standard deviation 48 pages). Seven Centres produced reports that were < 10 pages long. In these short reports reporting was perfunctory. Three Centres, however, each produced a report over 80 pages long, i.e. more than 1 standard deviation longer than the average length. These three reports were also the reports with extensive participant quotes. The common factor among these three reports was that they all came from Centres where the research team included specialist qualitative researchers. There was no other obvious pattern to the variation in report detail (such as country, language or research team experience). It is thus possible that the detail of reports was related to the composition and research approach of individual research teams.

These differences in reporting were probably compounded by the different levels of specificity within the guidelines, with some reporting tasks being outlined more comprehensively than others. While careful consideration was given to the size of groups, participants' backgrounds, the venue and choice of moderator, there was less instruction on other aspects of the study. For example, the requirements for reporting the focus groups findings ran to just two paragraphs of the guidelines. Although it may have been assumed that experienced researchers would write consistent reports, this was not the case, which reflects the practical realities of conducting qualitative health research. The implication is that the more quantitative tasks were easier for Centres to report, whereas sifting through transcriptions to produce qualitative data was more resource intensive and may have

required a specialist skill set that some Centres may have not have had. These reasons may explain why the translation and reporting of focus group results was less well reported than some administrative tasks. Since these reports formed the basis on which future instrument development work proceeded, there are implications regarding ecological validity.

A third source of variation may relate to collecting data across cultures, particularly in developing countries where there may be some strategic methodological and logistic challenges including different team expectations, problems of language and communication, culturally sensitive interpretation of tasks, access to participants, availability of technology and a lack of resources and research funding [34, 35]. These issues may in part explain the differences in guideline interpretation reported by some Centres. Finally, although every effort was made to ensure regular bi-annual meetings and constant contact of the principal investigators through email, the geographical distance between Centres meant that the investigators proceeded independently. Thus, although standardization of data collection was partially achieved through the use of the guidelines, this was subject to local interpretation and therefore local variations in data collection occurred.

The implications for future multi-centre cross-cultural studies are that protocols should be fulsome and complete and that full specification of the results to be reported must be included. Other helpful suggestions arising from this study would include protocols that are published, shorter tasks for older adults or others who may be participating in studies at times of personal illness (e.g. the use of shorter QoL instruments), teams with appropriate skill composition, common researcher training and monitoring of local researchers by a single research team. It would appear that despite a growing use of focus groups for the development and evaluation of instruments in various cultures, comprehensive sets of procedures or requirements for the international part of development and evaluation may be lacking in refinement.

Overall this study highlights the need for the formulation of clear, detailed, well designed and practical protocols to encourage consistent approaches by field centres and enable in-depth comparison across international sites. The rationale for each of the requirements should also be evident to each member of the research team. This is particularly the case when protocols are open to cross-cultural interpretation.

Conclusion

This paper reports the level of adherence to a set of international guidelines developed for use by WHOQOL Field Centres to conduct and report focus groups aimed at eliciting key concepts of QoL among older adults.

The findings show that there was general consistency in carrying out and reporting of the agreed guidelines. Where there were inconsistencies or poor reporting, Centres were asked the reasons for this. In general, the reasons given by Centres for these discrepancies indicated a combination of a lack of specificness in reporting requirements of the guidelines and structural issues around *ad hoc* local difficulties. The more detailed reports were those from Centres with qualitative researchers on the research teams. These findings are consistent with the literature, and suggest that multi-centre cross-cultural protocols need to be thoroughly developed, that research teams need to have the appropriate blend of skills, there needs to be appropriate researcher training in implementation procedures, and that high standards of reporting or auditing are required if results are to be comparable.

Acknowledgements

The authors would like to thank Professor Mick Power for his leadership of the WHOQOL-OLD project, Ms Rowena McElhinney and Ms Rosemary Taylor for their support of the WHOQOL Field Centres during data collection.Ethics approvals were gained through local ethics committees for each of the WHOQOL Field Centres. The WHOQOL-OLD study was funded through the European Commission, Contract number QLRT-2000-00320. This paper was prepared through funding from a Melbourne University International Collaboration grant and support from the Australian Victorian Public Health Trainee Scheme.

References

1. McNeil WH. Plagues and Peoples. New York: Anchor Press, 1998.
2. Skevington SM. Advancing cross-cultural research on quality of life: observations drawn from the WHOQOL development. World Health Organisation Quality of Life Assessment. Qual Life Res 2002; 11: 135–144.
3. de Haes JC, Olschewski M. Quality of life assessment in a cross-cultural context: use of the Rotterdam Symptom Checklist in a multinational randomised trial comparing CMF and Zoladex (Goserlin) treatment in early breast cancer. Ann Oncol 1998; 9: 745–750.
4. 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: 1417–1432.
5. Sartorius N, Kuyken W. Translation of health status instruments. In: Orley K, Kuyken W (eds), Quality of Life Assessment: International Perspectives. Berlin: Springer-Verlag, 1994: 3–18.
6. Anderson RT, Aaronson NK, Bullinger M, McBee WL. A review of the progress towards developing health-related quality-of-life instruments for international clinical studies and outcomes research. Pharmacoeconomics 1996; 10: 336–355.
7. 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: 451–459.
8. Fox-Rushby J. The relationship between health economics and health-related quality of life. In: Orley J, Kuyken W (eds), Quality of Life Assessment: International Perspectives. Berlin: Springer-Verlag, 1994: 61–74.
9. Hawthorne G, Richardson J. Measuring the value of program outcomes: a review of utility measures. Exp Rev Pharmacoeconomics Outcomes Res 2001; 1: 215–228.
10. Wagner AK, Gandek B, Aaronson NK, Acquadro C, Alonso J, Apolone G, et al. Cross-cultural comparisons of the content of SF-36 translations across 10 countries: Results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998; 51: 925–932.
11. ABS. National Health Survey: SF-36 Population Norms, Australia. ABS Catalogue No. 4399.0 ed. Canberra: Australia Bureau of Statistics, 1995.
12. Pike KL. Language in Relation to a Unified Theory of the Structure of Human Behaviour. The Hague: Mouton, 1966.
13. Orley J, Kuyken W. Quality of Life Assessment: International Perspectives. Berlin: Springer-Verlag, 1994.
14. WHOQoL Group. Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL). Qual Life Res 1993; 2: 153–159.
15. Fayers PM, Hopwood P, Harvey A, Girling DJ, Machin D, Stephens R. Quality of life assessment in clinical trials – guidelines and a checklist for protocol writers: the U.K. Medical Research Council experience. MRC Cancer Trials Office. Eur J Cancer 1997; 33: 20–8.
16. Olschewski M. Compliance with QOL assessment in multicentre German breast cancer trials. Stat Med 1998; 17: 571–575.
17. WHOQoL Group. The World Health Organization Quality of Life Assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995; 41: 1403–1409.
18. WHOQoL Group. Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment. Psychol Med 1998; 28: 551–558.
19. WHOQoL Group. The World Health Organization Quality of Life Assessment (WHOQOL): Development and general psychometric properties. Soc Sci Med 1998; 46: 1569–1585.
20. WHOQoL Group. WHOQOL-Bref: Introduction, Administration, Scoring and Generic Version of the Assessment. Geneva: World Health Organization, 1996.
21. Stewart DW, Shamdasani PN. Focus Groups: Theory and Practice. Newbury Park, Sage: 1990.
22. Ferrell BR, Grant MM, Funk B, Otis-Green S, Garcia N. Quality of life in breast cancer survivours as identified by focus groups. Psycho-Oncology 1997; 6: 13–23.
23. Lydick E, Zimmerman SI, Yawn B, Love B, Kleerekoper M, Ross P, et al. Development and validation of a discriminative quality of life questionnaire for osteoporosis (the OPTQoL). J Bone Miner Res 1997; 12: 456–463.
24. Bass EB, Jenckes MW, Fink NE, Cagney KA, Wu AW, Sadler JH, et al. Use of focus groups to identify concerns about dialysis. Choice Study Med Decis Making 1999; 19: 287–295.
25. Cottrell CK, Drew JB, Waller SE, Holroyd KA, Brose JA, O'Donnell FJ. Perceptions and needs of patients with migraine: a focus group study. J Family Practice 2002; 51: 142–147.
26. Shen W, Kotsanos JG, Huster WJ, Mathias SD, Andrejasich CM, Patrick DL. Development and validation of the Diabetes Quality of Life Clinical Trial Questionnaire. Med Care 1997; 37: AS45–66.
27. WHOQOL Group. The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In: Orley J, Kuyken W (eds), Quality of Life Assessment in Health Care Settings. Berlin: Springer-Verlag, 1994: 41–60.
28. Lucasey B. Qualitative research and focus group methodology. Orthop Nurs 2000; 19: 54–55.
29. Freeman K, O'Dell C, Meola C. Focus group methodology for patients, parents, and siblings. J Pediatr Oncol Nurs 2001; 18: 276–286.
30. Microsoft. Microsoft Excel 2000. Seattle: Microsoft, 1999.
31. QSR International. NVivo. 2.0 Melbourne: QSR International, 2002.
32. SPSS. SPSS for Windows. 10.0 Chicago: SPSS Inc., 2000.
33. GraphPad. InStat. 3.02 San Diego: GraphPad Software, 2000.
34. Mapanga K, Ndlovu R, Mapanga M, Mudokwenyu-Rawdon C, Brooten D, Morris D, et al. A model for international research collaboration. Int Nurs Rev 1999; 46: 117–121.
35. Kearney N, Miller M, Sermeus W, Hoy D, Vanhaecht K. Multicentre research and the WISECARE experience. Workflow Information Systems for European Nursing Care. J Adv Nursing 2000; 32: 999–1007.

This study was funded by the European Commission Fifth Framework, QLRT-2000-00320, and was carried out under the auspices of the World Health Organization Quality of Life Group (WHOQOL Group).

The WHOQOL-Old Group comprises a coordinating group and collaborating investigators in each of the following field centres: Professor M. Power, Dr K. Quinn, K. Laidlaw, University of Edinburgh, UK; Dr R. Lucas, Insititut Català de l'Envelliment, Barcelona, Spain; Professor S. Skevington, F. McCrate, University of Bath, UK; Dr M. Amir*, Y. Ben Ya'acov, T. Narkiss, Department of Behavioural Science and Social Work, Ben-Gurion University of the Negev, Israel; Professor L. Kullman, National Institut for Medical Rehabilitation, Hungary; G. Bech-Anderson, Psychiatric Research Unit, Copenhagen, Denmark; Professor Ji-Qian Fang, Sun Yat-sen University of Medical Sciences, Guangzhou, China; Professor M.C. Angermeyer, I. Winkler, Department of Psychiatry, University of Leipzig, Germany; A/Professor G. Hawthorne, Ms N. Davidson, Department of Psychiatry, University of Melbourne, Australia; Dr M. Kalfoss, Faculty of Nursing, Menighetssosterhjemmets College, Oslo, Norway; Dr A. Leplege, INSERM, Paris, France; Dr E. Dragomirecka. Prague Psychiatric Centre, Prague, Czech Republic; Dr M. Martin, Health Services Research Associates Inc, Seattle, U.S.A.; Dr M. Tazaki, Department of Science, Science University of Tokyo, Tokyo, Japan; Professor M. Eiseman, Department of Psychology, University of Tromsoe, Norway; B. Nygren, Department of Nursing, Umea University, Sweden; Dr A. Molzahn, Faculty of Human and Social Development, University of Victoria, Canada; Dr J. Ceremnych, Scientific Department of Gerontology Problems, Institute of Experimental and Clinical Medicine, Vilnius, Lithuania; Dr M. Fleck, Department de Psiquiatra e Medicina Legal, Porte Alegre, Brazil; Professor K-F. Leung, Hong Kong Hospital Authority, Kowloon, Hong Kong; Dr N. Von Steinbuchel, Centre of Neurogerontolopsychology, Geneva University of Psychogeriatrics, Geneva, Switzerland; Assoc Professor E Eiser, Celal Bayar University, Manisa, Turkey; Dr L. Schwartzmann, Dept of Medical Psychology, Uruguay; Dr R. Kilian, Department of Psychiatry, University of Ulm, Germany.

Address for Correspondence: Graeme Hawthorne, Department of Psychiatry, The University of Melbourne, Parkville, Victoria 3010, Australia
Phone: +61-3-8344-4000; Fax: +61-3-9349-2792
E-mail: graemeeh@unimelb.edu.au

*We note with great sadness the death of Dr. Marianne Amir in January 2004.