

The use of qualitative methods in developing the descriptive systems of preference based measures of health related quality of life for use in economic evaluation

Katherine Stevens¹, Simon Palfreyman²

¹Health Economics and Decision Science (HEDS), ScHARR, The University of Sheffield, UK.

²Sheffield Teaching Hospitals NHS Foundation Trust, UK.

Name and address for correspondence:

Katherine Stevens

Health Economics and Decision Science

School of Health and Related Research, University of Sheffield,

Regent Court, 30 Regent Street

Sheffield S1 4DA, UK

Telephone: 0114 222 0841

Fax: 0114 272 4095

Email: K.Stevens@Sheffield.ac.uk

Abstract

The objective of this paper is to describe how qualitative methods may be used in the development of the descriptive systems of preference based measures (PBM) of health related quality of life.

The requirements of NICE and other agencies together with increasing use of patient reported outcome measures (PROMS) has lead to an increase in the demand for PBM. Recently, interest has grown in developing new PBM and whilst previous research on PBM has mainly focused on the methods of valuation, research into the methods of developing descriptive systems is an emerging field.

Traditionally, descriptive systems of PBMs were developed using top down methods, where the content was derived from existing measures, the literature, or health surveys. A contrasting approach is a bottom up methodology, which takes the views of patients or lay people on how their life is affected by their health. This approach generally requires the use of qualitative methods. Qualitative methods lend themselves well to the development of PBMs. They also ensure the measure has appropriate language, content validity and responsiveness to change. Whilst the use of qualitative methods in the development of non PBMs is fairly standard, their use in developing PBMs was until recently nonexistent.

In this paper, we illustrate the use of qualitative methods by presenting two case studies of recently developed PBMs; one generic and one condition specific. We outline the stages involved, discuss the strengths and weaknesses of the approach and compare with the top down approach used in the majority of PBMs to date.

Background

The need to develop rigorous methods for effectively describing the impact of healthcare interventions on patients has been driven by two factors. The first is the desire to collect data on the quality of care from the patient's rather than the clinician's perspective. The second is the need to compare interventions, through economic evaluation, in order to use limited health care resources more efficiently. Both of these make use of questionnaires that seek to obtain information from the patient regarding their health related quality of life, and are often referred to as Patient Reported Outcome Measures (PROMS).

PROMs use within the UK National Health Service (NHS) was stimulated by a key recommendation of the Darzi report, "High Quality Care for all"ⁱ published under the last Labour Government. This specified that the impact of treatments on quality of life should be measured through the routine use of PROMs. As a result, since 2009 NHS providers have been required to ask patients to complete a PROM before and after four surgical procedures (hip replacements, knee replacements, hernia repair, and varicose veins). These data are now published on a monthly basis. The General Election in 2009 resulted in a change of Government but did not decrease the importance of PROMs within the NHS. The Coalition Government's White Paper: Liberating the NHS: Transparency in outcomes - a framework for the NHSⁱⁱ, also stressed the importance of PROMs and advocated their more widespread use.

The use of economic evaluation as a tool to aid health care resource allocation decision making has increased markedly in recent years and has been formalised in many countries through agencies such as the National Institute for Health and Clinical Excellence (NICE) in the UKⁱⁱⁱ and others around the world^{iv,v,vi}. These decision making bodies require evidence on the cost effectiveness of health care interventions under consideration as part of the decision making process. The majority of these agencies have formal guidelines for the methods of economic evaluation and whilst in the past these guidelines have not stipulated the measure of benefit for cost effectiveness analysis, more recently they have explicitly stated that health effects be measured in quality adjusted life years (QALYs)^{iv, iiiiv}. In 2004 NICE introduced its reference case (the set of methods considered most appropriate by NICE)

and stated QALYs as the measure of benefit. All submissions to NICE now require a cost effectiveness analysis based on these methods with health effects measured in QALYsⁱⁱⁱ. The QALY combines length of life and quality of life into a single summary measure. As well as being able to take account of changes in quality of life, quantity of life or both, it is useful in health care resource allocation decision making as it allows comparison across clinical areas due to the use of a common measure of benefit^{vii}.

QALYs are calculated as the product of the time spent in a particular health state multiplied by the utility or preference weight associated with that health state. This preference weight is on a scale where 1 is perfect health and 0 is equivalent to being dead (with negative values for health states judged to be worse than dead). These preference weights for health states can be obtained in a number of different ways, including the use of expert opinion, literature, direct valuation from the patient or the use of preference based measures (PBM)^{vii}. PBM are a type of PROM that typically consist of a health state classification system (HSCS) and a set of preference weights for each of the health states defined by the HSCS. Usually, patients complete the HSCS which defines their current health state and then the preference weight assigned to that health state can be used to calculate QALYs.

There are two main types of PBM; generic and condition specific. A generic PBM is intended to cover all areas of health and should be applicable to any clinical condition. An example is the EQ5D which has been widely used in numerous clinical conditions^{viii}. A condition specific PBM is only concerned with a particular condition, for example asthma or diabetes. They may be used when there are concerns that a generic PBM may not be valid or reliable, for example by not being sensitive enough. Examples include the SQOL 3D for sexual quality of life^{ix} and the AQL 5D for asthma^x.

Generic PBMs such as the EQ-5D, SF-6D, HUI2 and HUI3 have sometimes been found to be inappropriate or insensitive for some conditions^{vii}. If a generic measure is not regarded as appropriate, then a condition specific PBM could be developed which would be relevant to the condition. The most common method to date to do this has been to modify an existing condition specific non PBM^{xi}

however an alternative is to develop a new measure from scratch. Other reasons for developing a new PBM are for specific populations such as children, adolescents or the elderly where it is felt that the existing measures (mainly adult generics) are not appropriate for these populations.

The main constraint of developing a PBM is that the health states defined by the HSCS should be amenable to valuation. Health state valuation is the way in which the preference weights for the health states are determined. It can be undertaken using a variety of methods including the standard gamble, time trade or ordinal methods such as ranking and discrete choice experiments.^{vii} To be amenable to health state valuation, HSCS should have dimensions with ordinal levels and ideally one item per dimension. In addition, there is a limit to the number of dimensions that it can contain. Typically people can value 7 (plus or minus 2) pieces of information at any one time^{xii} and so the number of dimensions is typically no more than 9. The most widely used generic descriptive systems range from 5 to 9 dimensions^{vii}. This is a practical constraint on the number of dimensions within a descriptive system as it is unlikely that respondents would be able to handle a larger number when undertaking valuation exercises. Non PBMs of HRQoL do not have to operate within these constraints and hence can have much larger descriptive systems.

The requirements of NICE and other agencies together with increasing use of PROMS in the NHS has lead to an increase in the use of PBM in health care research and as a consequence, an increase in the demand for PBM. In the past, attention has been focused on the methods of health state valuation and less on the methods of developing the HSCS. Recently however, there has been increasing interest in developing new PBMs, particularly condition specific ones^{xi} and so there is increasing interest in the methods of development.

The main existing generic PBM for adults have all used a top down approach in the development of their descriptive systems, that is, the content has been derived from existing literature, instruments and health surveys. The Measurement and Valuation of Health Survey which was used to develop the EQ-5D, used 196 members of the general population to validate five existing descriptive systems, by

surveying lay concepts^{xiii}. The Quality of Well Being (QWB) drew its items mainly from an existing US Health Interview Survey and Social Security Administration Survey^{xiv} the SF-36 (from which the SF-6D is derived) used data from existing instruments^{xv} and the Assessment of Quality of Life (AQOL) was developed from a literature review from 1970 and interviews and focus groups with 24 clinicians^{xvi}. The HUI2 was developed from a review of epidemiological surveys and a review of the literature which generated a large pool of potential attributes. They then used a sample of child and parent pairs to rate these items, to select attributes for inclusion. The Health Utilities Index 3 (HUI3) was developed from the HUI2 by increasing the number of dimensions to 8 (through the separating out of some dimensions and the removal of others) and increasing the number of levels for all dimensions to between 5 and 6^{xvii}.

Three more recently developed generic PBMs for children are the AQoLY^{xviii}, the EQ5DY^{xix} and the CHU9D^{xx,xxi,xxii}. Both the AQoLY and the EQ5DY were developed using top down methods, as they have been adapted from the existing adult versions of their measures (AQoL and EQ5D respectively). In contrast, the CHU9D was developed using bottom up qualitative methods and is one of the case studies outlined in this paper.

In a recent systematic review of condition specific PBM, Rowen^{xi} found that just over half (12 out of 22) of condition specific PBM identified were derived from a single non PBM for the condition of interest, often using a combination of factor, psychometric and Rasch analyses to determine a HSCS that is amenable to health state valuation. The other 10 had developed “de novo” classification systems, of which 7 used a top down approach of taking items and/or dimensions from existing questionnaires, the literature, expert opinion, using psychometric techniques and some limited qualitative research in combination with this. The other 3 measures were based on qualitative research and unfortunately little detail was given about the methodology. They all took a similar approach of interviewing patients using semi structured or unstructured interviews and then an expert panel selected domains from these interviews and created levels for these. None of these measures were developed to be QALY measures intended for health care resource allocation, their stated purpose was

as a clinically useful outcome measure. All three follow the same valuation method and have anchors on the 0 worst state and 100 best state scale, not consistent with the QALY model. The review found that overall many studies poorly described the methodology used to develop the measures, particularly the development of the classification system. This lack of detail makes it difficult to develop methodology in this area and the review recommends that further research examining best practice and providing recommendations for the development of CSPBM should be undertaken.^{xi}

A contrasting approach to the top down methods outlined is a bottom up methodology, which takes the views of patients or lay people, seeking their input on how their quality of life is affected by their health problem or condition. This approach generally requires the use of qualitative methods to generate the items or content of the descriptive system, either through the use of focus groups or individual interviews. **Error! Bookmark not defined..** The advantages of this method are that the final measure developed will have appropriate language and terminology for the population which should increase the content validity. It will also improve responsiveness to change, as it will ensure outcomes of relevance to the patient are included. **Error! Bookmark not defined..** A further benefit of this approach is that it ties in with the NHS priority of putting the patient at the centre of decision making, providing patient centred care and public-patient involvement in health care and research. In more recent years, the importance of involving patients and lay people in the development of all types of QoL measures has been more widely recognised^{xxiii} and whilst the use of qualitative methods in the development of non PBMs is now fairly standard, their use in developing PBMs for use in economic evaluation was until recently nonexistent.

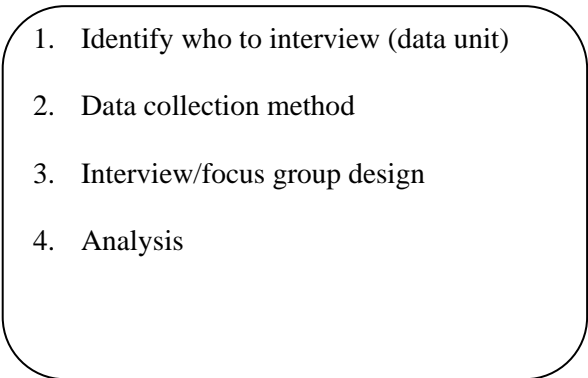
The first application of using a bottom up methodology with qualitative methods to develop a measure intended for use in economic evaluation (although not as a QALY measure) was undertaken by Grewal et al^{xxiv} in their development of a QoL measure for older people. In this work, they interviewed people over 65 years of age to determine attributes for a new index focusing on QoL for older people. They conducted in depth interviews with 40 participants to explore their views about what was important to them in terms of QoL. From these interview data they directly developed five

conceptual attributes. This approach of determining attributes or dimensions directly is in contrast to the more common approach taken in the general QoL instrument development literature, where large lists of items are generated from interviews, the literature and/or opinion and then a technique such as factor analysis is used to develop the dimensions^{xxv, xxvi}.

There are a number of critical stages in qualitative research design^{xxvii}. The first of which is to review your research question and decide what type of qualitative data you require. For the development of PBM, generated data is required, which is data generated by respondents who give their own interpretation and explanation from interviews, as opposed to naturally occurring data that you may observe in a natural setting. This is appropriate because you need to generate the data by asking individuals about how their health affects them, this is not something that would be naturally occurring and you could observe. Secondly your data unit and time span need to be considered, for example case studies of individuals and whether at a point in time (cross sectional) or over a series of time points. The development of a PBM most likely demands cross sectional data as you are interested in views across the spectrum and not how people's views change over time. Thirdly a data collection method needs to be determined, for example focus groups or individual interviews. Fourthly, you need to devise your sampling strategy and finally an analytical approach (e.g. thematic content analysis) needs to be determined, which will largely depend upon the purpose of the research, and can be aided with a data management tool such as NVIVO.

These stages of qualitative research design can be applied to the development of the descriptive systems of PBM and require a series of key choices to be made in the stages of development. These are summarized in Figure 1 below and then outlined in more detail.

Figure 1: The key stages of using qualitative methods to develop a descriptive system for a PBM

- 
1. Identify who to interview (data unit)
 2. Data collection method
 3. Interview/focus group design
 4. Analysis

5. Development of the descriptive system

1. Identify who to interview

Firstly, a decision has to be made about the relevant population(s) to interview. Perhaps the most relevant population, particularly for a condition specific PBM is the patient population who have the condition, as they have first hand knowledge about how it affects their life. Other relevant populations could include carers, family members, medical professionals or other experts. One important reason for focusing on the patient population is that others can either overestimate or underestimate the impact of a disease on quality of life^{xxviii}. For example, clinicians tend to be more generic and less focused, and to underestimate the social and subjective aspects of a disease compared to patients^{xxix}. Once the decision regarding who to include is made, thought needs to be given to the sampling frame used. If interviewing patients, they could be sampled purposively on the basis of their health, so as to include as wide a range as possible. Other sampling criteria could include age, gender and ethnicity to ensure that a wide range of views are incorporated.

2. Data collection method

There is a choice to be made between using focus groups, individual interviews or a combination of the two. Individual interviews offer more sensitivity and depth whereas focus groups allow participants to feed off each others ideas and can stimulate more discussion. In a focus group participants can feel more comfortable, however if the material being discussed is sensitive, it may be better to do individual interviews as they give the opportunity for complete privacy. Focus groups may lead to inhibitions in raising issues that some may feel are important, yet others do not and so they are unwilling to share them. However, individual interviews can feel uncomfortable for shy people and some may feel nervous. Thomas et al^{xxx} found that there was no difference in terms of the depth of data generated between focus groups and interviews. Interviews and focus groups can also be used to complement each other and ensure a breadth of data are obtained. For example focus groups could provide the initial analysis which can then be explored in more depth by individual interviews or focus groups could be used to validate data obtained from interviews^{xxxi,xxxii}. Ultimately the choice between these two methods of data collection needs to be decided with reference to the population

being considered and the practical advantages and disadvantages of each method **Error! Bookmark not defined.**

3. Interview/focus group design

An important consideration for the development of PBMs is whether to bring existing material to the interview or focus group, for example using prompts or suggestions from the literature or other sources, or to have a blank canvas. The advantage of using prompts is that they can help to stimulate ideas and make use of the existing knowledge base and evidence. The advantage of having a blank canvas is that there is no bias and no pre conceived ideas are introduced. Effective interviewing is a complex skill and interviewers need the appropriate training. Unstructured interviews can be particularly difficult to conduct and is a skill that needs practice and reflection^{xxxiii}. Semi-structured interviews generally use a topic guide and provide the interviewer with an interview schedule and prompts in order to direct the interviewee but allow them to explore other areas where appropriate.

xxxiv

4. Analysis

Before analysing the data, it is important to check whether data saturation has been achieved, that is, when no new data is emerging from the interviews or focus groups^{xxvii}. It is important to achieve saturation for the development of PBMs in order to make sure that all areas of health related quality of life are captured by the measure.

When analysing the data, there are two general approaches that can be taken. Firstly, items can be identified from the interviews and then dimensions developed from these, perhaps through the use of factor or Rasch analysis. This mirrors the approach taken in the general QoL literature for instrument development.^{xxxv} Alternatively, the use of qualitative methods allows for the identification of dimensions directly from the data, such as that referred to earlier by Grewal^{xxiv}.

5. Development of the descriptive system

Lastly, once dimensions have been identified, levels need to be developed in order to form a HSCS amenable to valuation. The first decision is whether the levels should be based on the frequency or severity of a particular dimension. For example a dimension about pain could be either about the frequency of the pain or the severity of the pain. This choice may depend on the nature of the condition if developing a condition specific PBM or may be related to the purpose of the instrument. Qualitative data could be used to inform this choice by examining the way in which the dimensions or items were described in the interview and/or focus groups. Once this decision is made, wording for the levels needs to be developed. Existing scales, such as those from existing instruments or standard Likert type scales could be used or an alternative approach would be to develop new ones using the qualitative data.

Case Studies to illustrate the use of qualitative techniques in the development of PBMs.

Recent research by Stevens^{xx} and Palfreyman^{xxxvi, xxxvii} took a bottom up methodology to the development of new PBMs using qualitative methods. We present both of these as case studies in the context of the key stages outlined above. The Child Health Utility 9D (CHU9D) is a preference based generic paediatric health related quality of life measure^{xx}. The SPVU-5D is a condition-specific preference based instrument that was developed for use with venous leg ulceration.

Stage 1: Identify who to interview

CHU9D - Potential relevant populations to consider included parents/guardians, paediatric health care professionals, the general population, health care decision makers and children. All these were considered and children were chosen as the population to interview. This was because they were seen as the most relevant as they were the population the measure was intended for. By interviewing children, content validity would be increased **Error! Bookmark not defined.** and the language and terminology would be more appropriate than using the views of others. In addition, there is evidence that parents views are affected by their own health status, knowledge, experience and expectations^{xxxviii}. There may also be areas or contexts of the child's life that are unknown or less

well known to parents such as school^{xxxix}. Drotar^{xl} also notes that many children are capable of recognizing and appreciating aspects of their own health, such as symptoms and the impact of changes in their health which their parents may not notice. Purposive sampling was used with a primary criteria of age, followed by level of health (as rated by the parent) to ensure that the views of a full range of health across the ages were covered^{xx}. Secondary criteria were gender and ethnicity.

SPVU-5D - Both patient and professional perspectives were sought in developing the instrument. Although patients were central to the identification of the items for inclusion it was felt important to include the views of health professionals. This was in order to check that no important aspects were omitted and to further validate the items that were identified by the patients. A final consideration was that if health professionals were potentially to use the final instrument they had to be stakeholders in its development. For patients, purposive sampling was used in order to ensure that participants recruited approximated the range of ages, sex, and ethnicity attending the leg ulcer clinics in Sheffield. Purposive sampling was also used to ensure that the full range of clinicians involved in the care of venous ulcer patients were included. The sample included medical and nursing staff from both primary and secondary care involved in the care.

Both the CHU9D and SPVU-5D instruments used a purposive sampling approach in order to include the breadth of experiences of health of their respective populations. The aim was not to achieve statistical representativeness, but to capture the experience of health through identifying groups who possessed characteristics relevant to the health conditions being studied^{xxvii}.

Stage 2: Data collection method

CHU9D - In choosing between focus groups and individual interviews several considerations pertinent to children were relevant. The advantage of focus groups is that children can feel more comfortable and feed off each other's ideas, however the material being discussed was sensitive. In addition, pre existing dynamics within the groups of children selected can have an influence.

Individual interviews give the opportunity for complete privacy and the opportunity for children to

raise any issues that they wish. Individual interviews were chosen primarily because of the sensitivity of the topic and also because of the practicalities of managing young children in interview/groups situations; it is easier to manage an individual interview than a focus group of young children.

SPVU-5D – Both interviews and focus groups were used to collect data. The aim was for these to provide different sources of information and perspectives. One benefit of focus groups is the added data obtained through the interaction of the members. Using two different methods of data collection had the benefit of providing validation of the data interpretation from the semi-structured interviews and as a means to ensure that data saturation had been achieved.

Contrasting approaches were taken in developing the two instruments due to the nature of the populations interviewed and type of instrument being developed. It would not have been realistic to expect young children to participate in focus groups. Another issue was imposed by the nature of the instruments being developed. The CHUD9D being a generic measure meant that interviewees may have raised a range of health problems. However, the SPVU5D was a condition specific measure so respondents were more likely to feel comfortable in a focus group as they all shared the same experience of one particular health problem/illness.

Stage3: Interview/focus group design

CHU9D - A semi structured interview format was used, using a topic guide to ensure that each interview covered the same general questions (asking the child about their health problems and the impact on their lives). The topic guide did not include any items from the literature or existing instruments to use as prompts or discussion points as it was felt that these may have been developed for a different purpose and/or may not be the views of children and as there is a danger of acquiescence bias with children, including pre existing data may lead or bias the direction of the interview and what the children think. The interviews were therefore not influenced by preconceptions about what should be included. As this measure was developed with the intention of being child centred, it was felt that including data from other view points would conflict with this.

SPVU-5D - A semi-structured interview format was used but the interview schedule was generated from a review of the literature which highlighted specific symptoms and areas for exploration. The issues and areas that arose in earlier interviews, which gave insights into the impact of venous ulceration on quality of life, influenced the focus and discussion of subsequent interviews. A number of participants were also re-interviewed and were presented with the data analysis in order to check that it was a valid interpretation of the interviews.

Although the basis of the interview schedule for the two instruments differed, both techniques are equally valid. Using the literature to highlight areas for exploration can have the benefit of including a wider experience of the health problem and can also ensure that the final instrument incorporates any shortcomings of existing measures. The alternative of allowing the participants to direct the interviews has the advantage of grounding the instrument more firmly within their experience. Care has to be taken with this approach to ensure that data saturation has been reached and that no key areas have been missed.

Stage 4: Analysis

As the CHU9D was being developed specifically to be a PBM, the approach to the analysis was to identify dimensions directly from the data and then to develop levels for each of these as this suits the structure needed for a PBM. Thematic content analysis was used as it suits the approach of generating dimensions from this type of interview data. The analysis was guided by the research question; how does health affect children's lives, and the aim was to identify dimensions of their HRQoL^{xx}.

SPVU-5D – A thematic analysis of the data was undertaken using Framework analysis techniques^{xli}. It is particularly suitable for studies within public health and where there are clear aims at the outset – in this case identification of items for inclusion in a quality of life instrument^{xliii}. The interpretation of the data was examined through re-interviewing a number of participants and through using a focus group as method of triangulation. This method aimed to ensure that the analysis had resonance with the experience of venous ulcer patients and that data saturation had been achieved.

Both instruments sought to ensure that the qualitative data was “trustworthy” and “dependable”^{xliii},^{xliv}. This aims to demonstrate that the interpretation of the data was grounded in the data and not the researcher’s preconceived ideas. In addition, the SPVU-5D included triangulation and member checking (or respondent validation). Triangulation in qualitative research is essentially the use of different methods and perspectives in order to move away from a reliance on a single source of data so allowing the researcher to check the breadth of the data collection and the verification of the interpretation of the data^{xxvii}.

Stage 5: Development of the descriptive system

CHU9D - The qualitative data was used to inform whether the dimensions would be frequency or severity based^{xxi}, by examining the ways in which the children described the dimensions. This has the advantage of ensuring that the resulting descriptive system accurately reflects how children think about and describe their HRQoL. This led to the dimensions being severity based which also suits a descriptive system with a short recall period which is desirable for a paediatric measure^{xlv}. The qualitative data was again used to develop levels for all the dimensions as there were no suitable existing scales in the paediatric literature^{xxi}. The advantage of using the qualitative data in this way is that the terminology of the children is retained in the measure and the validity and practicality of the measure should therefore be increased.

SPVU-5D – The qualitative data was used to identify items to be incorporated into the new instrument. For some items it was clear from the qualitative interviews that severity and frequency were an issue. People were concerned with how severe and frequently pain occurred or how offensive and frequently their legs smelled. For other items it was how often they were able to undertake a specific activity or the frequency with which a problem impacted on their quality of life rather than the severity of the impact. A choice was made to have five levels within the dimensions in order to impose a limit on the potential number of health states. The wording for the levels was based on a

review of the current instruments and intended to reflect the same ordering as the other instruments but also aiming to ensure clear separation between the item levels. A recall period of seven days was chosen based on the treatment regime for venous ulceration which is based around compression bandages being usually being left in place for a week. The wording of the questions and the ordering of the levels of the dimensions was assessed by piloting the questionnaire on a convenience sample of clinicians and patients.

The development of the descriptive system of both measures was informed by the qualitative data and the approach to the analysis. In the case of the CHU9D, this led to dimensions being developed directly and their type (frequency or severity) and scales being informed by the qualitative data. In the case of the SPVU-5D, the qualitative data again informed whether the dimensions of the instrument should be severity or frequency based. For some of the items, it was clear from the qualitative interviews that both severity and frequency were an issue, e.g. people were concerned with how severe and frequently pain occurred.

Discussion

Although both the case studies outlined above used qualitative methods in the development of their descriptive systems, they differed in a number of ways with regard to the choices made at each stage in the development of the instruments. This was often dictated by the type of measure and the population it was intended for, in that for the condition specific measure it was more likely to increase the richness of the data collection from the use of focus groups as the disease condition was quite specific. Within the focus group patients can “feed off” each others experience. Whereas for the generic measure the aim was to cover a range of clinical conditions. Both measures used a similar approach to sampling, in order to achieve a breadth of experience. They also used qualitative data to develop the content of the instrument and to inform the development of the levels of the HSCS.

A bottom up methodology using qualitative methods offers several advantages over the traditional top down approach. By involving patients and/or lay people, the HSCS will contain appropriate language

and terminology and the dimensions will be of relevance to the population being studied, ensuring greater content and face validity. In addition, the measure will be more patient centred and as it will capture outcomes of relevance to patients, this means it is more likely to be responsive to change^{xlvi}. Taking a bottom up methodology ensures that the measure is developed according to its purpose, rather than relying on information from previous research that may have had different aims and objectives.

Measures developed using bottom up methodology are more amenable to self completion by the patient which is desirable as it is increasingly recognized in clinical trials and health services research that descriptions of the experience of a health state should be elicited from the patients in order to reflect the actual experience of the disease and its treatment^{xlvii}. As the CHU9D was developed using bottom up methodology, it is much more likely that children will be able to self complete as the language and terminology was all determined by children. If a child can provide reliable and valid data, then self report is the optimal.^{xlviii} Similarly, the SPVU-5D should be very amenable to self completion by the patient.

Qualitative approaches can be used to either generate dimensions directly or to develop a long list of items which then requires further quantitative work to classify these items into dimensions. Both of these differing approaches have been illustrated in this paper. The key difference in the approach of generating dimensions directly is that it is perhaps easier to develop a PBM since it generates one item per dimension and levels within these in order to construct a HSCS. This is the ideal for health state valuation. In contrast, the item generation approach has a stronger use and reliance on psychometric techniques to develop the dimensions.

Conclusions

In this paper we have outlined the importance of using qualitative methods in developing descriptive systems for PBM used in economic evaluations. We have described the stages involved and illustrated these through the use of two case studies.

The recent increase in the number of PBMs means that more attention needs to be paid to the methods used in their development. Particular attention needs to be focused on the development of the descriptive systems for PBMs, an area which has received very little attention in the past. The use of a bottom up methodology offers several advantages in this respect over the use of the more traditional top down approach used by the majority of existing PBMs. Content and face validity should be increased and the measure will contain more appropriate language and terminology. In addition, the dimensions will be of relevance to the population being studied and the measure will be more patient centred and amenable to self completion. Using this bottom-up methodology means that PBMs can be developed for purpose rather than having to rely on data from existing instruments. These may have been developed for other purposes rather than for use in health care resource allocation decision making. The use of qualitative techniques means that PBMs will be more likely to have a valid and reliable descriptive system. This should be the standard for any PROM and further research in this area for PBM is required.

References

-
- ⁱ Darzi 2008 High Quality Care for all
- ⁱⁱ Liberating the NHS: Transparency in outcomes - a framework for the NHS (Dept. of Health 2010)
- ⁱⁱⁱ National Institute for Clinical Excellence. Guide to the Methods of Technology Appraisal. April 2004. NICE.
- ^{iv} PBAC (Pharmaceutical Benefits Advisory Committee).
<http://www.health.gov.au/internet/main/publishing.nsf/Content/Pharmaceutical+Benefits+Advisory+Committee-1> Accessed 8th February 2011
- ^v CADTH (Canadian Agency for Drugs and Technologies in Health).
<http://cadth.ca/index.php/en/publication/35> Accessed 8th February 2011
- ^{vi} <http://www.pharmac.govt.nz/> Accessed 8th February 2011
- ^{vii} Brazier, J.E, Ratcliffe, J., Salomon, J. & Tsuchiya A. 2007. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford University Press.
- ^{viii} <http://www.euroqol.org/> Accessed 21st April 2011
- ^{ix} Ratcliffe, J., Brazier, J., Tsuchiya, A., et al. Using DCE and ranking data to estimate cardinal values for health states for deriving a preference-based single index from the sexual quality of life questionnaire. *Health Economics* 18, 1261-1276 (2009)
- ^x Yang Y, Brazier J, Tsuchiya A, Young T. Estimating a Preference-Based Index for a 5-Dimensional Health State Classification for Asthma Derived From the Asthma Quality of Life Questionnaire *Medical Decision Making* (2010) [forthcoming].
- ^{xi} **CosMEq study – reference to come**
- ^{xii} Miller GA. The magical number seven plus or minus two: some limits on our capacity to process information. *Psychological Review*. 1956. 63; 81-97
- ^{xiii} Van Agt, H, Bonsel, G. In Kind P, Brooks R, Rabin R (eds). EQ-5D concepts and methods, a developmental history. Springer; 2005.
- ^{xiv} Kaplan, R.M. & Anderson, J.P. 1988. The Quality of Well-Being Scale: rationale for a single quality of life index. In: *Quality of life: Assessment and Application*. Walker, S. R. & Rosse, R. M. (Eds.). London, MTP Press.
- ^{xv} Short Form 36. <http://www.sf-36.org/tools/SF36.shtml#VERS2>. Accessed 13/03/2008
- ^{xvi} Hawthorne, G., Richardson, J., Osborne, R. & McNeil, H. 1997. The Assessment of Quality of Life (AQoL) instrument: construction, initial validation & utility scaling. Working Paper 76. CHE research papers series. Monash University, Australia.
- ^{xvii} Feeny, David, William Furlong, George W. Torrance, Charles H. Goldsmith, Zenlong Zhu, Sonja DePauw, Margaret Denton, and Michael Boyle, "Multi-Attribute and Single-Attribute Utility Functions for the Health Utilities Index Mark 3 System." *Medical Care*, Vol. 40,(2) 2002, pp 113-128.
- ^{xviii} Moodie M, Richardson J, Rankin B, Sinha K, Lezzi A. Predicting time trade-off health state valuations of adolescents in four pacific countries using the AQoL-6D instrument. Centre for Health Economics, Monash University Research Paper 2009 (43).
- ^{xix} Wille N, Badia X, Bonsel G et al. Development of the EQ-5DY: a child friendly version of the EQ-5D. *Quality of Life Research* 2010; 19: 875-886.
- ^{xx} Stevens, K J. Working With Children to Develop Dimensions for a Preference-Based, Generic, Pediatric Health-Related Quality-of-Life Measure. *Qualitative Health Research*. 2010; (20); 340 - 351
- ^{xxi} Stevens K J. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Quality of Life Research*. 2009; 18 (8): 1105-1113.
- ^{xxii} Stevens K J. Assessing the performance of a new generic measure of health related quality of life for children and refining it for use in health state valuation. *Applied Health Economics and Health Policy*. 2011; 9 (3): 157-169
- ^{xxiii} Guyatt, G., Mitchell, A., Irvine, E. J., Singer, J., Williams, N. Goodacre, R., & Tompkins, C. 1989. A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology*, 96(3), 804-810.
- ^{xxiv} I. Grewal, J. Lewis, T. Flynn, J. Brown, J. Bond, J. Coast. Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science & Medicine*, Volume 62, Issue 8, Pages 1891-1901

-
- ^{xxv} Drotar, D. 2004. Validating measures of paediatric health status, functional status, and health related quality of life: key methodological challenges and strategies. *Ambulatory Pediatrics*. 4(suppl). 358-364.
- ^{xxvi} Riley, A. W., (2004). Evidence that school-age children can self report on their health. *Ambulatory Pediatrics*, 4(4), 371-376.
- ^{xxvii} Ritchie, L., Lewis, J. (Eds.) 2005. *Qualitative Research Practice; A Guide for Social Science Students and Researchers*. SAGE publications, London.
- ^{xxviii} Stein, R.E. K. 2004. Measurement of Children's Health. *Ambulatory Pediatrics*. 4, 365-370.
- ^{xxix} Ammerman, D. J., Watters, J. M. D., Clinch, J., Herbert, P. C., Wilson, K. G., Morriss, D. B., and Fergusson, D. Exploring quality of life for patients undergoing major surgery: The perspectives of surgeons, other healthcare professionals, and patients. *Surgery* 141[1], 100-09. 2007.
- ^{xxx} Thomas, L., McMillan, J., McColl, E., Hale, C. & Bond, S. 1995. Comparison of focus group and individual interview methodology in examining patient satisfaction with nursing care. *Social Sciences in Health*. 1, 206-220.
- ^{xxxi} Kitzinger, J. "Qualitative Research: Introducing focus groups." *BMJ* 311.7000 (1995): 299-302.
- ^{xxxii} Sullivan, J. M. and A. T. Foltz. "Focus groups: giving voice to people." *Outcomes Management for Nursing Practice* 4.4 (2000): 177-81.
- ^{xxxiii} Corbin, J. & Morse, J. M. 2003, "The Unstructured Interactive Interview: Issues of Reciprocity and Risks when Dealing with Sensitive Topics", *Qualitative Inquiry*, vol. 9, no. 3, pp. 335-354.
- ^{xxxiv} Barbour, R. S., Featherstone, V. A., & WoReN, M. o. 2000, "Acquiring qualitative skills for primary care research. Review and reflections on a three-stage workshop. Part 1: using interviews to generate data", *Family Practice*, vol. 17, no. 1, pp. 76-82.
- ^{xxxv} Fayers P, Hays R eds. *Assessing quality of life in clinical trials*. Second edition. Oxford University Press. 2007
- ^{xxxvi} Palfreyman, S. J. "Assessing the impact of venous ulceration on quality of life." *Nursing Times* 104.41 (2008): 34-7.
- ^{xxxvii} Palfreyman, S. J., J. A. Michaels, and J. Brazier. "Development of a tool to examine the effect of venous ulcers on patients' quality of life." *Nursing Standard* 21.45 (2007): 57-69.
- ^{xxxviii} Petrou, S. 2003. Methodological issues raised by preference-based approaches to measuring the health status of children. *Health Economics*. 12: 697-702.
- ^{xxxix} Matza, L. S., Swensen, A. R., Flood, E. M., Secnik, K. & Leidy, N. K. 2004. Assessment of Health Related Quality of Life in Children: A Review of Conceptual, Methodological, and Regulatory Issues. *Value in Health*. 7(1) 79-92.
- ^{xl} Drotar, D. 2004a. Measuring Child Health: Scientific Questions, Challenges, and Recommendations. *Ambulatory Paediatrics*. 4, 353-357.
- ^{xli} Snape, D. & Spencer, L. 2003, "The Foundations of Qualitative Research," in *Qualitative Research Practice. A guide for Social Sciences Students and Researchers*, J. Ritchie & J. Lewis, eds., Sage, London, pp. 1-23.
- ^{xlii} Ritchie, J. & Spencer, L. 1994, "Qualitative data analysis for applied policy research," in *Analyzing Qualitative Data*, A. Bryman & R. G. Burgess, eds., Routledge, London, pp. 173-194.
- ^{xliii} Glaser, B. G. & Strauss, A. 1967, *The discovery of grounded theory. Strategies for Qualitative Research*. Sociology Press, Chicago.
- ^{xliv} Lincoln, Y. & Guba, E. 1985a, "Establishing Trustworthiness.," in *Naturalistic Inquiry*, Sage, Beverley Hills, pp. 289-331.
- ^{xlv} Landgraf, J. M. 2005. Practical considerations in the measurement of HRQoL in child/adolescent clinical trials. In Fayers, P.& Hays, R. (Eds.), *Assessing Quality of Life in Clinical Trials*. 2nd edition. Oxford University Press
- ^{xlvi} Guyatt G H and Cook D J. (1994). Health status, quality of life, and the individual. *JAMA*, 272 (8); 630-631
- ^{xlvii} Fitzpatrick R, Davey C, Buxton M, Jones DR (1998). Evaluating patient based outcome measures for use in clinical trials. *Health Technology Assessment* 2 (14): 1-74
- ^{xlviii} Matza, L. S., Swensen, A. R., Flood, E. M., Secnik, K. & Leidy, N. K. 2004. Assessment of Health Related Quality of Life in Children: A Review of Conceptual, Methodological, and Regulatory Issues. *Value in Health*. 7(1) 79-92.