The Development of a Preference Based Paediatric Health Related Quality of Life Measure for use in Economic Evaluation

Katherine Josie Stevens
BSc (Hons) (York), MSc (York)

A thesis prepared for the degree of Doctor of Philosophy

School of Health and Related Research
The University of Sheffield

08 December 2008

Acknowledgements

I would like to thank my advisors on this research, Professors John Brazier and Christopher McCabe and my expert advisory panel:

Professor Sir David Hall (Professor of Community Paediatrics)
Professor Christine Eiser (Professor of Child Health Psychology)
Professor Mike Campbell (Professor of Medical Statistics)
Professor Jenny Roberts (Professor of Economics)
Dr Paul Sutcliffe (Research Fellow)
Dr Alicia O'Cathain (Senior Research Fellow)

Several key people and organisations made this research possible and I would like to thank them. Firstly, the schools involved in the research; Firs Hill Community Primary School and Hunter's Bar Junior School and all the staff, parents and children who took part. Jeremy Hamm at The Children and Young People's Directorate at Sheffield City Council helped with access to and selection of schools. The Clinical Research Facility at Sheffield Children's NHS Trust hosted the clinical study and I would particularly like to thank Wendy Swann for her support, the staff on the wards and the parents and children who took part.

Thanks also go to Natasha Buxton for transcription of the interviews, the Centre for Research and Evaluation at Sheffield Hallam University for undertaking the valuation interviews, the Medical Research Council for funding the research and to staff in Health Economics and Decision Science at The University of Sheffield for their support, especially Donna Rowen, Tracey Young, Yaling Yang, Jill Carlton and Simon Palfreyman.

Lastly I would like to acknowledge and thank the support of my husband and family.

The Development of a Preference Based Paediatric Health Related Quality of Life Measure for use in Economic Evaluation

Contents

Chapter 1: Introduction............1

Chapter 2: Background............3

Chapter 3: Literature review............18

Chapter 4: Key decisions in developing a descriptive system............71

Chapter 5: Developing the dimensions............92

Chapter 6a: Developing a draft descriptive system............143
Developing the wording for the levels within the dimensions

Chapter 6b: Developing a draft descriptive system............174
Determining the order of the levels within each dimension

Chapter 7: Testing the draft descriptive system in a general paediatric population............186

Chapter 8a: Testing the draft descriptive system in a clinical paediatric population............238

Chapter 8b: Refinement of the descriptive system............279

Chapter 9: Valuation study............292

Chapter 10: Overall discussion, conclusions and future research............365

List of Abbreviations used in the Thesis

CBA	cost benefit analysis
CEA	cost effectiveness analysis
CMA	cost minimisation analysis
CUA	cost utility analysis
FDA	Food and Drugs Administration
FE	fixed effects
HRQoL	health related quality of life
MAE	mean absolute error
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
PBM	preference based measure
PRO	patient reported outcome
QALY	quality adjusted life year
QoL	quality of life
RE	random effects
RMSE	root mean square error
SG	standard gamble
TTO	time trade off
WHO	World Health Organisation

Summary

Use of economic evaluation to aid decision making is widespread and has increased in the UK especially since the introduction of the National Institute for Health and Clinical Excellence (NICE). Cost utility analysis, a form of economic evaluation, allows comparison of interventions within and between disease areas by using outcome measures that combine length and quality of life into a single summary measure, conventionally the quality adjusted life year (QALY). Generic preference based health related quality of life (HRQoL) measures have been developed for adults for this purpose, but research in paediatric populations is more limited.

A review of the literature of generic paediatric quality of life measures showed that no preference based measure existed which used children to develop the descriptive system. There was also very little methodological guidance on key issues in the development of new measures and no evidence was found on whether children share similar HRQoL frameworks across age.

This thesis presents work which has developed a new preference based paediatric HRQoL measure designed for use in economic evaluation. The descriptive system was developed from interviews with over 70 children in order to determine what dimensions of HRQoL were included. These were then used as the basis for developing the descriptive system by undertaking further analysis and empirical fieldwork with children.

This descriptive system was piloted with children in schools and hospital and the results used to refine the descriptive system to be suitable for valuation.

Preference weights were obtained by valuing a sample of health states with the UK general population and then modelling to obtain values for all the health states defined by the descriptive system. The results demonstrated that it is feasible to value this descriptive system and the new measure is starting to be used in health care evaluation studies worldwide.

Chapter 1

Introduction

1.1 Aims and Objectives of the Research

Aim:

The aim of this research was to develop a preference based paediatric health related quality of life measure for use in economic evaluation.

Objectives:

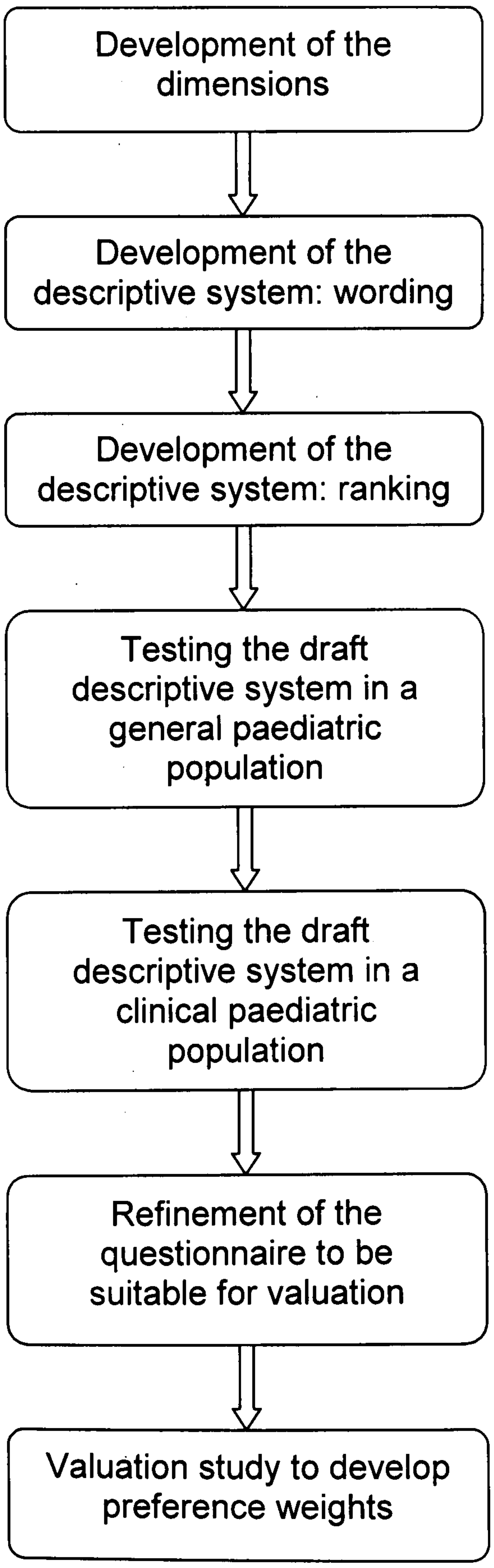
1. To review existing generic paediatric instruments.
2. To determine the relevant dimensions of paediatric health related quality of life.
3. To develop a descriptive system based on these dimensions which is amenable to health state valuation.
4. To evaluate the psychometric performance of the descriptive system.
5. To determine whether the descriptive system is suitable for valuation.

1.2 Structure of the thesis

Chapter 2 gives the background to the research, looking at the use of economic evaluation in resource allocation decisions, how benefit measurement is undertaken, the importance of measuring paediatric health related quality of life and why a new measure is needed. Chapter 3 is a literature review looking at existing paediatric generic measures and the evidence of their performance and assesses their purpose and suitability for use in economic evaluation. Chapter 4 considers the key decisions when developing a new health related quality of life measure, including the methods chosen. Chapter 5 describes how the initial dimensions were identified and chapters 6a and 6b describe how these were developed into a descriptive system. Chapters 7 and 8a describe the testing of the measure on general and clinical paediatric populations respectively. Chapter 8b describes how the descriptive system was refined to its final version to be suitable for health state valuation. Chapter 9 describes the valuation study which obtained preference weights for each health state defined by the descriptive system and Chapter 10 gives an overall discussion and conclusions from the work and considers future research in the area.

The flow chart below shows the key methodological stages of the work.

1.3 Flow chart of key methodological stages



Chapter 2

Background

2.1 Introduction

This chapter provides a background of health care resource allocation decision making in the UK where resources are limited and choices need to be made over which treatments or interventions to fund. The techniques of economic evaluation can be useful to aid this decision making and the different types are outlined here. The need for measuring health related quality of life is discussed and the need for research into outcome measurement in paediatric health related quality of life is considered.

2.2 Resource Allocation Decisions in the National Health Service (NHS)

The basic economic problem is that resources in the world are limited and scarce and there is a large demand for them. Therefore decisions have to be made about how to allocate these resources. The problem applies across all sectors, including education, defence and the environment and also applies in the field of health care as resources used to provide health care are not enough to meet all demands (Brazier 2007). By choosing to provide a health care intervention or service, an alternative use of those resources is sacrificed, which is known in economics as the opportunity cost (Drummond 1997).

The UK has a publicly funded health care system, the National Health Service (NHS) which is funded through general taxation and naturally has a limited budget with which to provide healthcare. There are many competing demands on this budget and new demands appear all the time as new treatments are developed or alternative ways of providing healthcare are introduced. Because these resources are limited, choices or decisions have to be made about how to allocate them. For example,

- Should a new drug be funded?
- Should a new piece of equipment be purchased?

- Should a breast cancer screening programme be offered to all women?

These types of decision have always had to be made, however, recently there has been a move towards making more informed, transparent and consistent decisions in health care, especially since the introduction of the National Institute for Health and Clinical Excellence (NICE) in England. NICE is an independent organisation which is responsible for providing national guidance on promoting good health and preventing and treating ill health (NICE).

2.3 Economic Evaluation

To help make these health resource allocation decisions in a consistent, informed and transparent manner, the techniques of economic evaluation can be used. Economic evaluation compares choices in terms of their costs and benefits, by looking at the ratio of the incremental cost (the difference in cost) compared to the incremental benefit (the difference in benefit). The two key features of economic evaluation are that it is comparative and both cost and benefits are considered (Drummond 1997).

Economic evaluation to aid health care resource allocation decision making in the UK NHS is being used and has increased over recent years especially since the introduction of NICE (Brazier 2007). In addition to England, there are other countries who take a similar approach to health care decision making and also advocate the use of economic evaluation, including the Pharmaceutical Benefits Advisory Committee in Australia (PBAC), and the Canadian Agency for Drugs and Technologies in Health (CADTH).

In order to undertake an economic evaluation both the cost and benefits of each alternative must be identified, measured and valued. Costs are usually reasonably straightforward however the benefits are often more complex. Health benefits in an economic evaluation can be measured in many ways, including clinical outcomes, such as the number of hip fractures avoided or amount of blood pressure reduction, in terms of quality of life, or in terms of length of life gained. The ratio of the incremental (i.e. extra compared to the alternative) cost and incremental benefit is called the incremental cost effectiveness ratio (ICER) and allows comparison of interventions within a disease area if the same outcome measure (or measure of benefit) is used, for example the number of asthmatic attacks avoided. However it is limited if you wish to compare across disease areas. This limited type of economic evaluation is generally termed cost effectiveness analysis (Drummond 1997).

There are other types of economic evaluation and the difference is the way the benefits are measured. Cost minimisation analysis can be undertaken if the benefits of 2 or more interventions are demonstrated to be equal and then it just becomes a case of minimising the cost. This is rarely the case in practice however, as benefits of different interventions are not usually equal and there is nearly always uncertainty around the estimate of benefits, meaning that you cannot be sure that the benefits are equal (Brazier 2007). In addition, there are very limited circumstances under which cost minimisation is appropriate. Briggs et al (Briggs 2001) argue that unless a study has been designed to show the equivalence of treatments (which is rare), it in not appropriate to undertake cost minimisation analysis on the basis of observed lack of significance of costs or benefits between treatments.

Another form of economic evaluation is cost benefit analysis where everything is valued in monetary terms, including the benefits. An intervention is then worthwhile if the benefits exceed the costs. Whilst this form of economic evaluation has several advantages, in that it allows comparisons of programmes across different sectors and can incorporate a much wider range of benefits, including non health benefits, in practice it is difficult to determine the monetary value for these benefits (Brazier 2007).

Cost consequence analysis is a form of economic evaluation which does not attempt to combine the costs and benefits into single numbers, but simply identifies them all and presents them to the decision maker, leaving the decision makers to decide about whether the intervention is worthwhile. This usually means however, that the reasons for making a decision become unclear and are less explicit than other forms of economic evaluation (Brazier 2007).

Finally, cost utility analysis, which can be seen as a special case of cost effectiveness analysis, compares interventions with benefits measured in terms of years in full health, where quantity and quality of life are combined into a single index. This has several advantages for decision making. Firstly, this is very relevant for health care, as it can capture the benefits of interventions that improve quantity of life alone, quality of life alone or a combination of the two. Secondly, it allows comparison of interventions across disease areas because the measure of benefit can be used in any disease area, giving a common denominator. This is very useful for decision makers who have to make decisions about competing healthcare interventions across the entire health sector. It can also be used to compare interventions with several health outcomes as the measure of benefit captures all these changes into a single number. It can also be used to compare interventions within disease areas that have different health outcomes as the benefits will be expressed in the same terms. The most common and widely used approach to this type of benefit measurement is the quality adjusted life year (QALY) (Drummond 1997).

2.4 Quality Adjusted Life Years (QALYs)

The QALY is a single index number which is generated from two components, length of life and quality of life. This is done by assigning a weight to the health state that a person is in and multiplying it by the time period they are in that health state. To calculate QALYs, this weight must be on a scale with a maximum value of 1 (equivalent to being in perfect health) with 0 being equivalent to death and values less than 0 possible, which indicates health states which are worse than being dead. For example, living in a health state with a quality adjustment weight of 0.7 for 5 years, would be 3.5 (0.7*5) QALYs, which is deemed equivalent to 3.5 years in perfect health. In reality, more complex health profiles are usually seen, with a person moving between different health states. In this case, the number of QALYs is calculated by summing the product of the value and time in each state.

The QALY is a useful index as it can capture changes in both quantity and quality of life. Measuring quantity of life is straightforward, however estimating the quality of life weight is more complex. It is essentially comprised of two components; describing the health state and valuing the health state.

There are a number of ways in which this can be done in practice. Existing off the shelf preference based measures (PBM) of health related quality of life (HRQoL) can be used, which have a defined health state classification system and an existing set of preference weights for each health state defined by the system. These measures can be either generic or condition specific and are sometimes termed indirect measures (Feeny 2005). Alternatively, bespoke vignettes can be developed which describe the condition and values can be elicited for these. The most commonly used methods to elicit these values include choice based methods such as the time trade off (TTO) or standard gamble (SG) techniques or the visual analogue scale (a non choice based technique) (Brazier 2007). The standard gamble technique is based on expected utility theory and asks respondents to make a choice between a certain outcome of the health state to be valued and an uncertain option with 2 possible outcomes, usually perfect health and dead. The probability of these possible outcomes is varied until the respondent is indifferent between the choices (Brazier 2007). The time trade off method again asks respondents to make a choice, however this time there is no risk involved, respondents are asked to trade off quality of life against length of life. The choice is between a set period of time (t) in the health state to be valued and a period of time x (where x<t) in full health. The value of the health state is determined from the point where individuals are indifferent between the 2 choices (Brazier 2007). The visual analogue methods asks respondents to rate health states on a scale with anchor points (often best health state and dead) such that the relative distance between locations reflects the difference between health states (Brazier 2007).

Values can also be elicited directly from patients, which avoids having to describe the state of health as patients just think about what they are currently experiencing. Again, values can be elicited using the techniques described above. This method is sometimes termed a direct measure (Feeny 2005).

In obtaining preference weights for the health states, we are concerned with the values attributed to the different components (dimensions) of a HRQoL measure (Kind 2005). It is this that will give cardinal values and quantify the difference between health states, meaning that this can be measured in a meaningful and useful way.

Finally, mapping from non preference based measures onto generic preference based measures can be undertaken, using regression methods to estimate a relationship between the measures and then applying this to a dataset. This approach is often useful when preference based data have not been collected in a study (Brazier 2007).

The QALY model has several assumptions associated with it. It is assumed that individuals will experience any number of different health states for different periods of time in different sequences throughout the course of their lives. If the QALY is to represent individual preferences over time, we need to make a number of assumptions about preferences. Firstly, the health state value must be independent of the duration of the health state, for example the value would be the same whether the health state lasted for 5 or 20 years. This is termed mutual utility independence, meaning that quantity and quality of life are independent.

Secondly, there is an assumption of constant proportional time trade off which means that if an individual is indifferent between 2 alternatives with 2 different health states for different lengths of time, the individual should remain indifferent between them when the duration is changed, as long as the durations are changed in proportion to one another.

Thirdly, there is an assumption that there is risk neutrality over life years. An example of this is given in Chapter 3 by Brazier (Brazier 2007):
An individual should be indifferent between the following 2 alternatives:

- Health state h for 2 years for certain
- A lottery with a 50-50 chance to live in health state h for either 1 year or for 3 years.

This is because the expected duration of the 2 alternatives is the same.

Finally, the health state value must be independent of when it occurs in a person's life, i.e. assuming that a health state will have the same value whenever the person experiences it and the health state value must be independent of what states occur before and/or after it. That is to say that the value is the same regardless of whether a person experiences a more severe or less severe state before it. This is known as additive separability. (Brazier 2007)

The above conditions are based on expected utility theory and are restrictive assumptions and there are several empirical studies which have demonstrated that individual preferences violate these conditions in a non systematic way (Tsuchiya and Dolan 2005).

To overcome these problems, one approach suggested has been to define the measure of health benefit as the number of years in full health equivalent to the expected health outcome (which takes account of the profile of health experienced over a persons life). This approach is called the health year equivalent (HYE) and there has been substantial debate in the literature about this, as it does not rely on the QALY model conditions between length and quality of life. However, the HYE is difficult to operationalize in practice. Other alternatives include adapting utility theory so that it can accommodate these violations in individual behaviour or accept that the violations occur but continue to use the theory as the normative basis on which to build policy decisions. (Brazier 2007)

Cost effectiveness analysis using the cost per QALY approach is much more useful for decision making as it allows comparison across all disease areas as benefits are in a common metric, hence the ICER is the incremental cost per QALY gained. The QALY approach is also recommended by NICE in its recommendations for the methods of economic evaluation (NICE 2004). The advantages of the QALY are that it generates a single index, it reflects the strength of preference for different outcomes and it incorporates both quality and quantity of life into a single measure.

2.4.1 Generic measures

Generic PBM are widely used in economic evaluation. They are generally easy to use and have the advantage that they are comparable across all disease areas and so are relevant to all patient groups. They are also broad in their coverage, attempting to capture all areas of HRQoL. Off the shelf measures such as the EQ-5D are straightforward to use as they consist of a simple 5 dimensional descriptive system and a set of existing preference weights that can be applied (Dolan 1997). They are easy to use in clinical trials as they are generally short and simple for patients to complete and the preference weights can be applied by the researcher. Use of a generic measure has also been recommended for use in the reference case analysis by NICE (NICE 2004).

2.4.2 Generic vs. condition specific measures

Generic measures are designed to be suitable for use in any population in any disease area, whereas condition specific measures are designed for use in specific populations, for example asthma or diabetes. One of the main advantages of condition specific measures is that they are more likely to be sensitive to changes and therefore will be better at detecting treatment effects (Brazier 2007). However, as they focus on a particular condition, the descriptive systems may exclude co-morbidities and complications which may alter the impact of the condition because of interactions between dimensions. Therefore the full extent of the impact on a patient's HRQoL is not accurately captured and important impacts may be missed. In addition, there may be focusing effects due to respondents focusing on the health condition or disease described by the measure and not thinking about other areas of HRQoL. For non PBM, there is also the problem of not being able to compare across disease areas, something which would not be an issue with using a generic measure. However, for PBM this can be overcome, because as long as the valuation of a descriptive system is consistent, then this puts measures onto a common metric.

2.4.3 Existing preference based measures

There are generic PBM that are widely used for adults, including the EQ-5D, the SF-6D, (Brazier 2002) the Health Utilities Index 3 (HUI3), the Quality of Well Being (QWB) and the AQoL. (Brazier 2007) They have all been used in research studies such as clinical trials for measuring the benefits of health care interventions and can be used to calculate QALYs so that cost utility analysis can be undertaken. The EQ-5D is also recommended in the guidance by the National Institute for Health and Clinical Excellence for reference case analysis (NICE 2004).

There is a current lack of generic preference based measures in children however and research in the paediatric field is much less. In his PhD thesis, McCabe reviewed existing measures of paediatric quality of life and assessed their suitability for use in economic evaluation (McCabe 2003). He concluded that measures for paediatric populations are lacking, research in this area is extremely limited, and that there is a need for a " robust, methodologically sound paediatric generic health related quality of life measure, which incorporates preferences across health states" and that "none of the existing measures fulfil this need in their present form". (McCabe 2003) This is discussed further in Chapter 3.

One approach to overcoming this gap would be to use adult measures, however they are not really suitable for use in paediatric populations as the dimensions may not be appropriate or pertinent to children and the response scales, wording and format may also not be appropriate. (Eiser 2001)

Recently there has been work to adapt the EQ-5D for children (Hennessy 2002). Primary school age children were sampled to test out the wording of the EQ-5D. The sample included children with experience of long standing illness. They found that there were some problems with understanding the wording used in the adult version of the EQ-5D and so altered the wording to become more child friendly and have produced a child friendly version. The preference weights are the same as for the adult version (Euroqol). The assumption made here is that the dimensions of health are the same for children as for adults, it is just the wording that needs adapting. This is a large assumption as there may be dimensions of HRQoL of relevance to children that are not included. Similarly there may be dimensions included that are of no relevance to children. This point is also noted by Landgraf (2005), in that adapting items from an adult instrument by rewording them to be relevant to children is not sufficient, because the instrument should be relevant at both the concept and the item level. In addition, Matza et al (2004) note that instruments designed to assess HRQoL among children with asthma should assess child functioning within the relevant contexts. They give the example of asthma where the impact on a child's life is likely to be different than that of an adult because of the context. For example limiting participation in play and activities would have social and emotional consequences different from those an adult may experience. They also note that because the impact of disease and treatment may be substantially different for adults and children, HRQoL outcomes from adult trials should not be applied to children, instead, children's HRQoL should be examined directly (Matza 2004).

Currently, the only PBM for children is the Health Utilities Index 2 (HUI2). The original version of the HUI (mark 1) was developed by Torrance et al specifically for use in an economic evaluation of neonatal intensive care (Torrance 1982). Since then, this has been followed by the development of the HUI2 and HUI3, of which the HUI2 was designed for use in children and the HUI3 in adults. The HUI2 originally consisted of 7 dimensions (sensation, cognition, mobility, emotion, self care, pain and fertility), each with between 3 and 5 levels and was first used with children with cancer, hence the inclusion of a fertility dimension (to capture the side effects of chemotherapy treatment) (Feeny 1992). By assuming the fertility dimension to be normal, the developers state that the HUI2 can be used as a generic measure (Torrance 1996). Whilst this measure has been widely used in practice, it is based on a 'within skin' concept of health, in that the dimensions are very symptom based rather than on the impact of the symptoms on the child's quality of life. It was also developed based on a review of the literature and incorporated the views of children, adults and the developers as to what was included in the final descriptive system. The preference weights were determined by parents of school age children.

2.5 Why measuring quality of life in children is important

It is important to measure HRQoL in children for a number of reasons. Firstly, recent medical advances have meant that there is an increase in child survival and consequently a focus on the quality of this life survived and not just the length. There are also now higher rates of survival of childhood conditions, often with the persistence of chronic health conditions /impairment (Stein 2004) and it is important to measure this.

Many interventions can have an impact beyond altering clinical variables (e.g blood pressure) and these important effects would not be detected unless HRQoL was measured too.

In addition, the ever increasing number of health care interventions means that there is more of a need to evaluate treatments in order to make decisions between competing alternatives. This requires that the benefits of interventions be measured.

There is also a movement in the NHS to focus more on the outcomes of treatment and clinicians are becoming more aware of effectiveness and outcome measurement.

The Food and Drug Administration (FDA) has issued regulations to encourage paediatric drug testing as part of drug development, however these trials have tended to focus only on efficacy and safety endpoints rather than looking at the impact on the children's HRQoL as well (Matza 2004). Measuring HRQoL can help to answer questions such as whether HRQoL has changed in patients over time and if so, in which direction and by how much.

There are different options to addressing this gap in the field of paediatric outcome measures and developing a new measure. One option would be to adapt an existing non preference based measure of health related quality of life for children into a preference based one. This approach has been undertaken before, for example in adults with the SF-36 (Brazier 2002) or in children with a disease specific measure, the ADQoL for children with atopic dermatitis (Stevens 2005). This depends whether there is a suitable existing measure in the paediatric field and is considered further in Chapter 3.

Alternatively, an entirely new measure could be developed. This would give the opportunity to develop the measure directly for its intended purpose, allowing the content and the structure to be determined early on.

2.6 Conclusion

Due to the growing importance of measuring quality of life in children, the increase in use of PBMs in economic evaluation for health care resource allocation decisions, and the lack of research in the important field of paediatric measures, there is a need to develop a new paediatric generic PBM of HRQol. The next chapter reviews existing paediatric measures to assess if there are any existing measures that could be adapted and if not, a measure will be developed from scratch.

This thesis firstly reviews the literature on existing paediatric generic PBM and assesses the suitability of these for use in economic evaluation (Chapter 3) and then reports on the development of a new paediatric generic PBM measure (Chapters 4-10), which aims to address some of the gaps in this area and improve on existing measures.

References

Brazier, J.E., Roberts, J. & Deverill, M. 2002. The estimation of a preference based measure of health from the SF-36. Journal of Health Economics. 21 (2), 271-292.

Brazier, J.E., Ratcliffe, J., Salomon, J. & Tsuchiya, A. 2007. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford University Press.

Briggs AH, O'Brien BJ. 2001. The death of cost-minimization analysis. Health Economics ;10 (2) :179-84.

CADTH (Canadian Agency for Drugs and technologies in Health).
http://cadth.ca/index.php/en/publication/35
Accessed 10/06/2008

Dolan, P. 1997. Modelling valuations for EuroQol Health States. Medical Care. 35, 1095-1108.

Drummond, M. F., O'Brien, B., Stoddart, G. L. & Torrance, G. W. 1997. Methods for the Economic Evaluation of Health Care Programmes. Oxford Medical Publications. Second Edition.

Eiser, C. & Morse, R. 2001. Quality-of-life measures in chronic diseases of childhood. Health Technology Assessment. 5(4).

Euroqol. http://www.euroqol.org/ accessed 16/06/2008

Feeny, D., Furlong, W., Barr, R. D., Torrance G. W., Rosenbaum, P. & Weitzman S. 1992. A comprehensive multi-attribute system for classifying the health status of survivors of childhood cancer. Journal of Clinical Oncology. 10(6), 923-928.

Feeny, D. 2005 Preference-based measures: utility and quality-adjusted life years. In Fayers, P. & Hays, R. (Ed.). Assessing Quality of Life in Clinical Trials. 2nd edition. Chapter 6.2. Oxford University Press.

Hennessy, S. & Kind, P. 2002. Measuring health status in children: developing and testing a child-friendly version of EQ-5D. Discussion Paper. 19th Plenary Meeting of the Euroqol group. York, UK, 13th-14th September.

Kind, P. 2005. Values and valuation in the measurement of HRQoL. In Fayers, P. & Hays, R. (Ed.). Assessing Quality of Life in Clinical Trials. 2nd edition. Chapter 6.1. Oxford University Press.

Landgraf, J. M., 2005. Practical considerations in the measurement of HRQoL in child/adolescent clinical trials. In Fayers, P. & Hays, R. (Ed.). Assessing Quality of Life in Clinical Trials. 2nd edition. Oxford University Press.

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

McCabe, C. May 2003. Estimating preference weights for a paediatric health state classification (HUI2) and a comparison of methods. Ph.D.Thesis. The University of Sheffield.

NICE (National Institute for Health and Clinical Excellence). http://www.nice.org.uk/ accessed 16/06/2008

NICE (National Institute for Clinical Excellence). April 2004. Guide to the Methods of Technology Appraisal.

PBAC (Pharmaceutical Benefits Advisory Committee).
http://www.health.gov.au/internet/main/publishing.nsf/Content/Pharmaceutical+Benefits+Advisory+Committee-1
Accessed 10/06/2008

Stein, R. E. K. 2004. Measurement of Children's Health. Ambulatory Pediatrics. 4, 365-370.

Stevens K J, Brazier J E, McKenna S P, Doward L C, Cork M J. The development of a preference-based measure of health in children with atopic dermatitis. British Journal of Dermatology,2005, 153:372-377

Torrance, G, W. 1982. Multiattribute Utility Theory as a Method of Measuring Social Preferences for Health States in Long-Term Care. In Kane, R. L. & Kane, R. A. eds. Values and Long-Term Care. Lexington, MA: Lexington Books Division of D. C. Health. 127-156.

Torrance, G. W., Feeny, D. H., Furlong, W. J., Barr, R. D., Zhang, Y. & Wang, Q. A. 1996. Multi-attribute utility function for a comprehensive health status classification system: Health Utilities Mark 2. Medical Care. 34(7), 702-722.

Tsuchiya A, Dolan P. 2005. The QALY model and individual preferences for health states and health profiles over time: a systematic review of the literature. Medical Decision Making. 25: 460-7.

Chapter 3

Literature Review

3.1 Introduction

This research will develop a new generic preference based paediatric health related quality of life (HRQoL) measure. The measure will be developed specifically for the purpose of making decisions about the allocation of paediatric health care resources and will be suitable for use in economic evaluations. Prior to developing this new measure, a literature review was carried out to assess the existing field of paediatric generic HRQoL measures and to inform the design of the new measure. This chapter reports on this literature review.

The purpose of assessing the existing field of measures was to consider their potential to be preference based measures (PBM), to check there are no PBM already existing for children and if found, to assess their performance in a paediatric population to date and more specifically, their suitability for use in economic evaluation.

To help inform the design of the new measure, the review also looked for evidence on whether children are able to provide information on their health related quality of life for the purpose of constructing or adapting a measure. This was to see if it is reliable and feasible to use children in the construction of a descriptive system. In addition, the review looked for any evidence that children share stable HRQoL frameworks across age in order to inform what age group the new measure was designed for.

As part of a large Health Technology Assessment (HTA) report, Eiser and Morse (2001) conducted a review of quality of life measures in children. They looked at many issues, including the extent to which adult measures are used on children, how appropriate this is, whether child and parent reports correspond and how feasible and reliable proxy measures of quality of life are in different disease contexts. Part of this report summarized paediatric generic quality of life measures identified by searching electronic databases from 1980 –July 1999, hand searching and searching the internet. Their search strategy was very general in order to identify any papers that were relevant to quality of life measures in children. They found 19 generic measures (and 24 disease specific measures) and summarized these in terms of respondent, age range, number of items, number of domains, and included evidence of their reliability or validity.

Their review was very comprehensive in that the search strategy was designed to be broad (which meant they had good sensitivity yet poor specificity) and included supplemental hand searching and internet searching. In addition, the coding of the papers was done by 2 independent researchers who cross checked for errors and omissions, increasing the validity of the review. The authors do acknowledge however that their search did focus on specific databases and they restricted their review to English language journals which may have lead to some bias in the inclusion of UK and US studies.

As part of his PhD thesis, McCabe (2003) updated the review by Eiser and Morse by searching the same databases until January 2002, including a further database and two review papers. The purpose was to review existing generic paediatric health related quality of life measures in terms of their suitability for use in economic evaluation. Generic measures were identified (most of these were included in the review by Eiser and Morse). He concluded that measures for paediatric populations are lacking and research in this area is extremely limited and that there is a need for a

> "robust, methodologically sound paediatric generic health related quality of life measure, which incorporates preferences across health states"

and that

> "none of the existing measures fulfil this need in their present form".

Whilst the review by MCabe was still comprehensive in that the search strategy was the same as that used by Eiser et al with the addition of extra databases, this review was much more specific, in that its purposes was to review existing paediatric measures in terms of their suitability for use in economic evaluation. All searching and reviewing was undertaken only by McCabe and so there is less validity than the Eiser review, however the review is clear and transparent in its methodology.

It has been a few years since these reviews and the field of paediatric health related quality of life has advanced since then, therefore this review will not only update but also extend the searches by including other relevant sources in the searching. In addition, further questions are asked of the literature in order to inform the design of the new measure and to understand the purpose of all existing instruments in the field. The aim was to address the following questions:

1a. What generic paediatric health related quality of life measures exist?
1b. What is their purpose and what are their key features?
1c. What is the evidence on their performance in a paediatric population?
1d. What is their suitability for use as a preference based measure and is there any potential to adapt the measure to be suitable for this purpose?

2. At what age is it feasible and reliable to elicit information about health related quality of life from children for the purposes of constructing a generic descriptive system?

3. Are there age groups which share a stable HRQoL framework?

3.2 Methods

As work has already been done in these previous reviews, this work builds on the review by McCabe (2003) who identified paediatric generic quality of life measures and reviewed their suitability for economic evaluation. The review covers the literature until January 2002. An update to this review was carried out in order to identify any new paediatric HRQoL measures developed since then. This was supplemented by undertaking additional searching on other well known databases to check for any paediatric HRQoL measures missed by the review by McCabe. In addition, all original references of the generic instruments included in the McCabe review were reviewed for evidence for questions 2 and 3.

Four separate parts of searching were therefore carried out:

1. An update to the review by McCabe (2003) using the same search strategy and databases.
2. Additional searching in quality of life databases (PROQLID and HSRR).
3. Additional searching in other databases not included in the McCabe review (DARE, NHSEED and HTA).
4. Checking the original references of all generic instruments included in the McCabe review.

1. Update to the review by McCabe

In his review, McCabe (2003) followed the search strategy of the review conducted by Eiser and Morse (2001). Eiser and Morse searched the following databases from January 1980 until July 1999:

1. MEDLINE,
2. BIDS ISI Science citation index
3. BIDS ISI Social Science citation index
4. PsychLIT
5. CCTR
6. Register of Controlled Trials

The following search strategy was used:

1. (quality of life) and (child* or adolesc*)
2. (health status or functional status or well-being) and (child* or adolesc*)
3. chronic illness or chronic disease or arthritis or asthma or cancer or cystic fibrosis or diabetes or epilepsy or AIDS or trauma or burns or technology dependent or low birthweight
4. 1 and 3

5. 2 and 3
6. 4 and (measure* or scale or index)
7. 5 and (measure* or scale or index)
8. self report or self-report or self assessment or self-assessment or child* report of adolesc* report
9. 4 and 8
10. 5 and 8
11. 1 and 8
12. 2 and 8
13. (parent or mother or carer) and (report or assessment)
14. 4 and 13
15. 5 and 13
16. 1 and 13
17. 2 and 13
18. (6 or 7 or 11 or 12 or 16 or 17) and (reliab* or valid*)

McCabe updated the review by Eiser and Morse by applying this strategy to the same databases for the period July 1999 to January 2002 and to further supplement the original search, the strategy was applied to the Econlit electronic database for January 1980 to January 2002. In addition, two further reviews were included (Landgraf and Abetz (1996), and Connolly and Johnson (1999).

Updated search for this review

Databases 1-7 listed below were searched using the same strategy as Eiser and Morse (and McCabe) for the period January 2002 – 13th December 2005. In addition, Embase was searched from 1980-13th December 2005 to further supplement the searching as this was not included in the previous reviews. Databases searched:

1. MEDLINE,
2. BIDS ISI Science citation
3. BIDS ISI Social Science citation index
4. PsychLIT
5. CCTR
6. Register of Controlled Trials
7. Econlit
8. Embase

Since the reviews by Eiser and McCabe were carried out, CCTR and register of controlled trials are now the same and are called CENTRAL and Psychlit is the same as Psychinfo. The Science and Social Science citation indexes were searched together.

Paper selection

Papers were reviewed to identify any new generic instruments, any papers with new evidence on the performance of the generic instruments and any papers with evidence for Q2 or Q3. As this review is only considering generic instruments, the following inclusion criteria were applied:

1. The measure is generic
2. The measure is designed to measure health related quality of life
3. The measure is designed for or has been used on children aged 18 years or younger.
4. It is a new instrument not previously identified or contains evidence for Q1c, Q2 and/or Q3.
5. The paper contains evidence from a measure that has been used in the English Language on an English speaking population.

The decision to only focus on generic measures was reached due to the very large volume of literature on disease specific instruments and the need to concentrate on generic instruments.

Papers were first reviewed at title and abstract stage and rejected or accepted and then the full paper reviewed if necessary.
Any review articles identified were used to cross reference and identify any further papers not identified by the searches.

2. Additional searching in PROQOLID and HSRR databases

In addition to the updated search, two other sources were searched for any generic measures and/or evidence for Q2 or Q3. The first was the PROQOLID database (Patient-Reported Outcome and Quality of Life Instruments Database (MAPI). This is a database of quality of life instruments and also contains evidence on their performance. It is managed by the MAPI research trust. The instruments are listed under categories and there is no specific search facility therefore all instruments listed under the categories paediatrics and adolescents were reviewed.

The second source was the Health Services and Sciences Research Resources (HSRR) Database (HSSR). This is a database of instruments/indices employed in health services research. There is no search facility, it is a list of all instruments, therefore all instruments in the database were reviewed. Both databases were reviewed on 23rd January 2006. The same selection and inclusion process as for part 1 was applied to the instruments identified.

3. Additional searching on DARE, NHSEED and HTA databases

A search was undertaken on the DARE (Database of Abstracts of Reviews of Effects), NHSEED (NHS Economics Evaluation Database) and HTA (Health Technology Assessment) databases on the Centre for Reviews and Dissemination (CRD) website at The University of York(CRD). As these are relatively small, specific databases, searching was broad so as not to miss anything. The search terms used were Child* and quality of life. This search was undertaken on 24th January 2006. The same selection and inclusion process as for part 1 was applied.

4. Checking all the original references of all generic instruments included.

All references for the development of the instruments identified (in the original McCabe review) were reviewed, to identify the purpose of the instrument (Q1b) and to see if they contained any evidence for Q2 and/or Q3.

3.2.1 Reviewing

Any new instruments identified were reviewed and details of the instrument included in an updated table of the instruments identified by McCabe (2003). This includes the instrument name, respondent, age-range suitable for, number of items and number of domains included in the McCabe review, plus an additional column stating the purpose of the instrument. In addition, any evidence on performance of the instruments not included by McCabe was reviewed and included in an updated table of the psychometric performance. This includes the same criteria used in the McCabe review: test-retest reliability; inter-rater reliability; internal consistency; face validity; content validity; construct validity; convergent validity and practicality.

Any papers identified to address Q2 and Q3 were reviewed for evidence.

Finally, an assessment was made of any new instruments identified in terms of their suitability for use in economic evaluation. This was done with the aid of the checklist developed by Brazier et al (1999). This checklist looks at practicality, reliability, validity and valuation methods (where applicable). It is reproduced in Appendix 3A.

There was only one reviewer of the literature (the author of this thesis). Ideally, there would be two reviewers for a review or systematic review to increase the validity, however project funding for this work only allowed for one reviewer. This review was not a full systematic review, it was a literature review.

3.3 Results

3.3.1 Updated search results

The table below shows the number of hits from each database and the total number of hits before and after the removal of duplicates.

Database	Hits
MEDLINE	199
BIDS ISI Science citation and BIDS ISI Social Science citation index combined	502
Psychinfo	33
Central	76

Embase	538
Econlit	0
Total including duplicates	**1348**
Total after removing duplicates	**975**

912 references were rejected after reviewing the title/abstract mainly because they were disease specific instruments or were not used on an English Language speaking population. This left 63 (975-912) references. A further 31 references were rejected after reviewing the full text. Of these 31 references, 4 were not generic measures, 6 were not designed to measure HRQoL, 1 had not been used on a population under 18, 7 contained no new evidence for questions 1c, 2 or 3, and 13 were not studies in English on an English speaking population. This resulted in 32 references (63-31).

The number of additional papers found from reviews was 3, therefore when added to the 32 references from above, the total number of papers accepted for review was 35.

3.3.2 Results from additional searching

Quality of Life Instruments Database:

10 instruments were identified in the paediatric instruments listing and 6 in the adolescent instruments listing. As some of these were same, this gave a total of 12 unique instruments. 3 were excluded as they had already been identified in the other searches, therefore 9 instruments were included for review, which were either new instruments or contained new psychometric evidence.

Health Services and Sciences Research Resources Database:

5 generic paediatric measures were found of which 4 met the inclusion criteria and were included for review.

CRD Databases (DARE, NHSEED and HTA)

42 hits for papers were identified, of which 1 was a duplicate with the update search already carried out. 31 were rejected at title/abstract stage and 8 were rejected after reviewing the full text. Therefore 2 papers were included for review.

Table 3.1 below is reproduced from McCabe (2003) and is updated to include new instruments found by the searches and updates to instruments already found by McCabe. These are highlighted by a shaded background.

Table 3.2 below is also reproduced from McCabe (2003) and contains details of the psychometric evidence for the instruments. New evidence found on instruments already identified by McCabe is highlighted in bold text. New instruments and their associated psychometric evidence are also highlighted in bold text.

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
Child Health and Illness Profile – Adolescent Edition	Child /adolescent	11 to 17 years	107 plus 46 optional disease/ injury specific items	6 (activity, comfort, perceived well-being, disorders, achievement, and resilience)	'The purpose of the instrument is to assess health in epidemiologic surveys, to determine the existence of systematic differences in health in subpopulations (including the socioeconomically disadvantaged), and to provide a basis for assessing the impact of changes in health services or health policies' (Starfield 1993)
Child Health Questionnaire	Child/**Proxy (MAPI)**	**5 to 18 years (MAPI)**	87 (youth form) **98 (parent form)** **50 (parent form) or 28 (short form) (MAPI)**	(CF87) 12 (physical functioning, role/social functioning, general health perceptions, bodily pain, role/social emotional, role/social behavioural, self-esteem, mental health, behaviour, family activities, family cohesion, change in health) **(PF50 and PF28) 14: general health, change in health, physical functioning, bodily pain/discomfort, limitations in school, work, and activities with friends due to physical problems and due to**	'The goal was to develop a comprehensive instrument that would be useful across a variety of healthcare settings and applications including academic research, clinical trials, physician offices, clinics, hospitals and health maintenance organizations' (Landgraf

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
				emotional and behavioural difficulties, behaviour, mental health, and self-esteem. emotional and time impact on the parent, limitations in family activities, and family cohesion (CHQ)	1999)
The Child Quality of Life Questionnaire	Child/Parent	9 to 15 years	45	15 (Getting about, doing things for self, soiling or wetting, school, out of school activities, friends, family relationships, discomfort due to bodily symptoms, worries, depression, seeing, communication, eating, sleep, appearance)	"assessment of quality of life in nine to fifteen year old children" (Graham 1997)
Dartmouth Picture and Word COOP charts	Child	12 years upwards	6	6 (physical fitness, emotional feelings, school work, social support, family communications, and health habits)	'To assess health and social problems of adolescents' For use as survey instruments and as a tool for detection of important problems. (Wasson 1994)
Exeter Health Related Quality of Life measure	Child	6 to 12 years	16 (reduced to 12 in a later version (each measured twice)	1 (health related quality of life)	'The EHRQL is designed to determine the impact of disease on everyday activities from the child's perspective and was constructed to assess

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					self-reported HRQL in children from the standpoint of perceived discrepancies between actual and preferred or ideal selves.' (Eiser 1999)
Functional Status II (R)	Parent	0 to 16 years <1 year; 1 year 2-3 years >= 4 years	43(long) 14 (short)	8 (communication, mobility, mood, energy, play, sleep, eating and toileting)	The original FS I was developed to measure individual child health status and characterize populations, the FS II is a revised version of this measure. This instrument was primarily designed to be a measure of the heath status of children with chronic physical conditions. (Stein 1990)
Generic Health Questionnaire	Child/parent	6 to 16 years (linguistically able children)	25	5 (general affect, peer relationships, attainments, relationship with parents, general satisfaction)	'To develop a measure suitable for assessing the quality of life for children with chronic illness.' (Collier 1997)
How are you?	Child/Parent	8 to 12 years	29	4 (physical activities, cognitive activities, social activities, and physical complaints)	'Developed in response to the need for a questionnaire that

Table 3.1

Instrument	**Respondent**	**Age-range**	**Number of items**	**Domains**	**Purpose**
					measures general as well as disease specific aspects of QoL of children with a chronic illness' The main purposes of HAY are to identify children who need additional care and to evaluate interventions. (Bruil 1997)
KINDL	Child	8 to 16 years	40	4 (Functional capacity everyday life, psychological well-being, physical state, social relationship)	Takes a psychometric approach. Generic instrument for quality of life assessment. (Ravens-Sieberer 1998)
Nordic Quality of Life Questionnaire for Children	Family	2 to 18 years	75	32 Global, external, interpersonal, personal: each with physical, mental, spiritual, social and economic interactions	
Paediatric Quality of Life Inventory	Child/Parent	8 to 18 years	45 (15 core, 30 supplemental)	3 core and 8 supplemental (Physical functioning, psychological functioning, social functioning, pain, nausea, procedural anxiety, treatment anxiety, worry, cognitive problems, perceived physical appearance, physician/nurse communication)	'is designed to be a generic paediatric quality of life measure to be utilized non categorically i.e. across various paediatric chronic health

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					conditions' 'to measure HRQoL outcomes for paediatric chronic health conditions' (Varni 1999)
Perceived Illness Experience	Child/Parent	7 years upwards	34 (in original). Subsequently revised to include a further 2 items on the physical appearance subscale and a new food subscale. No details of how many items this contains.	9 (interference with activity, disclosure of illness, school/work, peer rejection, parental behaviour, manipulation, pre-occupation with illness, food, physical appearance)	Originally developed to measure perceived illness experience in people with cancer, but may be used with other groups of children/young people with chronic illness. 'The study is an attempt to devise a method to assess the perceived impact of the illness from the child's point of view' (Eiser 1995) (Eiser 1999)
Quality of Life Profile – adolescent version	Adolescent	14 to 20 years	54	3 and 9 sub domains: Being (Physical, Psychological, Spiritual) Belonging (Physical, Social, Community) Becoming (Practical, Leisure, Growth)	'to develop a model and associated instrumentation to assess the quality of life of persons with

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					developmental disabilities' (Raphael 1996)
TAPQOL	Parent	1 to 5 years	43	4, 12 sub domains Physical Functioning (sleeping, appetite, lung problems, stomach problems, skin problems, motor functioning,) Social Functioning (problem behaviour, social functioning) Cognitive Functioning (communication) Emotional Functioning (positive mood, anxiety, liveliness)	'to measure parent's perceptions of HRQoL in preschool children and to evaluate the impact of diseases and treatments on the different domains of young children's lives' (MAPI 2006)
Warwick Child Health and Morbidity Profile	Parent	0 to 5 years	16 (ten primary and six subsidiary)	10 (general health, acute minor illness behavioural, accident, acute significant illness, hospital admission, immunization, chronic illness, functional health and health related quality of life)	'to provide a cross-sectional and longitudinal record of parentally reported health and morbidity of individual children and child populations' (Spencer 1996) 'has been designed to give a comprehensive picture of a child's health

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					and illness experience from the parental perspective' (Spencer 1996)
Health Utilities Index Mark 2	Child/Parent	6 to 16 (used in children down to 12 months in one study)	15	7 (Sensation, mobility, emotion, cognition, self-care, pain, fertility)	'to construct a utility or social preference based multi-attribute health and well-being index applicable to children aged 4-16 years' (Cadman 1986) 'HUI evolved in response to the need for a standardized system to measure health status and HRQL to describe: 1) the experience of patients undergoing therapy; 2) long-term outcomes associated with disease or therapy; 3) the efficacy, effectiveness and efficiency of healthcare interventions; and 4) the health status of general

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					populations.' (Health Utilities Index 2006)
Health Utilities Index Mark 3	Child Parent	6 to 16	15	8 (vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain)	'For measuring the overall health status and health related quality of life of individuals, clinical groups and general populations' (Furlong 1998)
16 Dimensions	Child	12 to 15 years	16	16 (mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, friends, physical appearance, mental function, discomfort and symptoms, depression, distress, and vitality	To develop a generic self assessment HRQoL measure for early adolescents. (Apajasalo 1996a)
17 Dimensions	Child	8 to 11 years	17	17 (mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, discomfort and symptoms, school and hobbies, friends, physical appearance, depression, anxiety, vitality, ability to concentrate, and learning ability and memory)	Generic measure of HRQoL for pre adolescents (Apajasalo 1996b)
Quality of Well Being	Parents/ adolescent children	4 to 18 years	3 plus 27 symptoms	3 (physical functioning, social/role functioning and mobility) plus 27 symptoms	'The purpose of the system is to express benefits and side effects of the program in terms

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					of equivalences of completely well-years of life' (Kaplan 1989)
TACQOL	Self (>8 years) Proxy (<8 years)	6-15 years	108 (parent form)	7 (parent form) (pain and symptoms, basic motor functioning, autonomy, cognitive functioning, social functioning, global positive emotional functioning, global negative emotional functioning)	'to develop a generic.. instrument to assess children's HRQoL' (Vogels 1998)
SF-10	Parent/guardian	5-18 years	10		Developed to address the need for scientifically valid health status assessment for the paediatric population (Health Services and Sciences Research Resources Database)
Child Health and Illness Profile- child version (CHIP-CE)	Self(child)/proxy(parent)	6-11 years	45 standard, 76 comprehensive parents report	5 (12 sub domains) for self (child) report (CRF) Satisfaction, comfort, resilience, risk avoidance, achievement. 6 (17 sub domains) for proxy (parent) report (PRF) All the above plus disorders.	To assess the health of children (Health Services and Sciences Research Resources Database) '..appears to be an effective tool for assessing children's perceptions of their own health and well-being that can be used to describe

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
					the health and functioning of populations of children. It is likely to be useful in monitoring outcomes of policies and services' (Riley 2004a)
KIDSCREEN	Self/proxy	8-18 years	3 versions: 52, 27, 10	10, 5, 1 Correspondingly 10 (Physical Well-Being, Psychological Well-Being, Moods and Emotions, Self-perception, Autonomy, Parent relation and Home Life, Social Support and Peers, School Environment, Social Acceptance (Bullying), Financial resources 5(Physical Well-Being, Psychological Well-Being, Parents and Autonomy, Social Support and Peers, School Environment) 1(General HRQoL Index)	Objective: 'Designed to obtain a self-report and an external assessment of health-related quality of life in children and adolescents in a wide range of ages and independently of current health status' (MAPI)
TedQL	Self/proxy	3-8 years	4 versions: 22, 23, 30, 22	4 versions correspondingly: 0 5 (Physical competence, peer acceptance, maternal acceptance, psychological functioning and cognitive functioning) 5 (Physical competence, peer acceptance, family acceptance, psychological functioning and cognitive functioning	to assess HRQoL in children aged 3-8 years (MAPI)

Table 3.1

Instrument	Respondent	Age-range	Number of items	Domains	Purpose
				0	
YQoL-R/YQoL-S	Self	11-18 years	2 versions: For group level perceptual: YQoL-S: (version for surveillance) 8 YQoLR: (version for program evaluation and research) 41	4 (self, relationship, environment, general quality of life)	'to assess multidimentionally the generic QoL of youth ages 11-18' (MAPI)
HSCS-PS	Clinician/parent	2.5-5 years	3-5 levels per dimension	12 (vision, hearing, speech, mobility, dexterity, self-care, emotion, learning and remembering, thinking and problem solving, pain and discomfort, general health, behaviour)	"To develop a multi-dimension health status classification system to describe pre-school children 2.5-5 years of age. (Saigal 2005)

Table 3.2

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
Child Health and Illness Profile – Adolescent Edition	r=0.49 to 0.87		α>0.7			√	√	
Child Health Questionnaire	**PF-50: ICC = 0.37-0.84** **Correlation from 0.17-0.74 (MAPI)** **ICC=0.49-0.78 (parent form at 2 weeks, with no significant event)** **ICC=0.05-0.82 (at 6 weeks with no significant**	**ICC=0.42-0.64 (Waters 2003)**	α>0.72 **PF:50** **α=0.65-0.96 (Asmussen 2000)** **α=0.79-0.92 (Waters 2001)** **α=0.07-0.86 (Georgalas 2004)** **α=0.75-0.97 (Wake 2003)** **α=0.60-0.93 (Waters 2000a)** **α=0.79-0.92**		√ **(Waters 2000b)** √ **(Waters 1999)**	√ **PF-50: mixed evidence (Asmussen 2000)** √ **(Williams 2005)** √ **(Geargalas 2004)** **√ (Waters 1999)**	√ **√ (Williams 2005)** **√ (Waters 2001)**	92% completeness **Time to complete: 8 minutes for parent form, 20 minutes child form (MAPI)** **CF-80: 70% response rate ((Waters 2001)** **Missing responses per item 0-18% (Geargalas 2004)**

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
	event) (Waters 2000a)		**(Waters 2001)** **α=0.19-0.87 (Waters 2000b)** **α=0.72-0.93 (CHQPF50, child 5-11 years)** **α=0.69-0.91 (CHQPF50, child 12-18 years)** **α=0.75-0.90 (CHQCF87, child complete 12-18 years) (Waters 1999)**			**√ (Waters 2000b)** **√ (Waters 2001)** **√ (Speechley 1999)**		**Mean time to complete = 13.1 minutes (Sung 2003)** **Response rate 70% (Waters 2001)** **Response rate 72% (Waters 2000b)** **Response rate 56.2% parents, 50.7% children (Waters 1999)**
The Child Quality of Life Questionnaire	r=0.4 to 0.7 for mothers & children	√		√ **(Graham 1997)**		√	√	Interviewer administered
Dartmouth Picture and	r=0.77		α=0.6 to 0.94	√		√	√	Less than 10 minutes to

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
Word COOP charts								administer
Exeter Health Related Quality of Life measure			α= 0.5 to 0.69		√		√	Interviewer administered. 20 minutes to complete
Functional Status II (R)	r=0.83 to 0.92		α=0.66 to 0.9		√	√		SF-14 used in a number of studies
General Health Questionnaire						√	√	25 item child complete questionnaire
How are you?	r=0.4 to 0.9					√ **√ (Le Coq (2000)**	√	**15 minutes for child to complete (Le Coq 2000) Evidence of responsiveness to change (Le Coq 2000)**
KINDL	r=0.8		α>0.7			√	√	High completion rates reported
Nordic Quality of Life Questionnaire								Postal questionnaire administered to

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
for Children								the whole family.
Paediatric Quality of Life Questionnaire		**0.45-0.50 between mother and child (Eiser 2003)** **0.358-0.635 between child self report and parent proxy (Uzark (2003)**	α=0.83 for child α=0.86 for parent **α=0.93 (parent)** **α=0.91 (child)** **(Varni (2002b)** **α=0.94 (mother)** **α=0.87 (child)** **(Eiser 2003)** **α=0.88 (child)** **α=0.90 (parent)** **(Uzark 2003)**			√ **√ (Varni 2002a)** **√ (Varni 2002b)** **√ (Uzark 2003)** **√ (Upton 2005)** **√ (Varni 2004)** **√ (Varni 2003)** **√ (Chan 2005)** **√ (Uzark 2003)**	√	**Missing data low (0.7% child, 3% parent) (Varni 2002b)** **Missing items 1.54% self report, 1.95% proxy report (Varni 2001)** **Missing data low (<3.5%) (Varni 2004)** **Item level missing – short form generic core scales 0-40% (Chan 2005)** **Missing data 1% (child) 2%**

Instrument	**Test-retest reliability**	**Inter-rater reliability**	**Internal consistency**	**Face Validity**	**Content Validity**	**Construct Validity**	**Convergent Validity**	**Practicality**
			$\alpha>0.70$ (Upton 2005) $\alpha=0.90$ (child) $\alpha=0.91$ (parent) (Varni 2004) $\alpha=0.88$ (child) $\alpha=0.89$ (parent) (Varni 2003) short form generic core scales: $\alpha=0.70-0.88$ (Chan 2005) $\alpha=0.75-0.91$ (parent proxy) $\alpha=0.70-0.81$					**(parent proxy) (Uzark 2003)**

Instrument	**Test-retest reliability**	**Inter-rater reliability**	**Internal consistency**	**Face Validity**	**Content Validity**	**Construct Validity**	**Convergent Validity**	**Practicality**
			(self report)					
Perceived Illness Experience			0.91 for total, 0.21 to 0.79 for sub-scales			√	√	Interview administered
Quality of Life Profile – adolescent version			α > 0.70 in all cases			√.	√	Poor, 40 minutes to complete with 'interviewer support'.
TAPQOL			α=0.43 to 0.84 **α=0.66-0.88 (MAPI)**		√	√	√	Response rates 60 to 88% **Time to complete: 10 minutes (MAPI)**
Warwick Child Health and Morbidity Profile	k=0.499 to 0.859	k=0.76-1.0				√		Completed by clinician. Maximum of ten minutes to complete.
Health Utilities Index Mark 2	Kappa = 0.02 to 1.0 r=0.55-1.0 **Kappa 0.59-**	r= 0.57-0.96	Not reported		√ **√ (Trudel 1998)**	√ **√ (Sung 2003)**	**√ (Trudel 1998)**	Short time to complete, self complete and proxy completion

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
	1.0 (Trudel 1998)					√ **(Speechley 1999)**		versions available. High completion rates **Mean time to complete = 3.1 minutes (HUI2/3) (Sung 2003)**
Health Utilities Index Mark 3	Question level 81.5 to 98.7% agreement; attribute 77.8 to 98.6%; provisional index r=0.767					√ **(Sung 2003)** √ **(Speechley 1999)**		Short time to complete, self complete and proxy completion versions available. **Mean time to complete = 3.1 minutes (HUI2/3) (Sung 2003)**
16 Dimensions	r=0.91					√		Short simple to complete questionnaire,

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
17 Dimensions	r=0.95					√		Interviewer administered.
Quality of Well Being	r>0.7	$\alpha>0.91$			√		√	15 minutes per interview
TACQOL			**PF: $\alpha=0.67$-0.84** **CF: $\alpha=0.65$-0.79 (MAPI)**			**√ (Vogels 1998)**		**Time to complete: 10 minutes (MAPI)**
SF-10								
Child Health and Illness Profile- child version (CHIP-CE)	**CRF: ICC = 0.63-0.76** **PRF: ICC = 0.71-0.85 (subdomains ICC=0.63-0.82) (MAPI)**		**CRF$\alpha=0.70$-0.82** **PRF$\alpha=0.79$-0.88** **$\alpha=0.68$-0.88 for subdomains (MAPI)**			**√ (CRF) (Riley 2004a)**		**Time to complete:** **CRF: 15 minutes (45 minutes in a group)** **PRF: 10-15 minutes** **PRF comprehensive: 20 minutes. (MAPI)** **Low rate of**

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
								missing data (Riley 2004a) Average of 21.4 minutes to complete for interviewer administration; 22.6 minutes for self completion (Riley 2004b) Response rate varied from 24%-78% (Riley 2004b) Low rate of missing data (Riley 2004b)
KIDSCREEN	Pearson's r 52 version: r=0.58-0.78 27 version: r=0.61=0.74		52 version: α =0.77-0.89 27 version: α =0.80-					Time to complete: 52 version: 15-20 minutes 27 version: 10-

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
	10 version: r=0.68 ICC 52 version: 0.56-0.77 27 version: 0.61-0.74 10 version: 0.55 (MAPI)		0.84 10 version: α =0.82 (MAPI)					15 minutes 10 version: 5 minutes (MAPI)
TedQL	ICC = 0.61-0.78 (MAPI)		α=0.53-0.76 (MAPI)					Time to complete = 15 minutes (MAPI)
Youth quality of life instrument – research version (YQoL-R) and version for surveillance (YQoL-S)	ICC 0.74-0.85 (Patrick 2002)		α=0.81-0.89 (MAPI) (Patrick 2002)			√ (Patrick 2002)		Time to complete: YQoL-R: 15 minutes YQoL-S: 2 minutes (MAPI)
Health Status Classification System for pre school	Kappa: 0.30-1 (Saigal 2005)	Kappa: 0.3-0.66 (Saigal 2005)				√ (Siagal 2005)	√ (Saigal 2005)	

Instrument	Test-retest reliability	Inter-rater reliability	Internal consistency	Face Validity	Content Validity	Construct Validity	Convergent Validity	Practicality
children (HSCS-PS)								

α = cronbach's alpha
r = Pearson correlation
ICC = intraclass correlation coefficient

What generic paediatric health related quality of life measures exist? (Question 1a)

There were 20 instruments identified in the review by McCabe (2003) which varied in terms of the age group they were suitable for, the number of items and dimensions, the perspective undertaken (quality of life or health related quality of life) and the respondent (child, parent, proxy or family). Since this review, there has been a number of paediatric generic quality of life measures developed which were found in the searches and in total 7 new instruments were added to the existing 20 instruments in the table (the TACQOL. SF-10, CHIP-CE, KIDSCREEN, TedQL, YQoL-R and the HSCS-PS).

The TACQOL is designed for children age 6-15 years and has a large number of items (108 for the parent form) which form 7 domains. Items were selected for inclusion based on psychological and clinical experience. It includes symptoms in its items, such as pain, headache, nausea, as well as basic motor, social, cognitive and emotional functioning. It was primarily designed to be completed by the parent about their child.

The SF-10, CHIP-CE, and the HSCH-PS were all derived from existing instruments (SF-36, CHIP-AE and HUI2/3 respectively). No evidence was found on the performance of the SF-10 and no information was found about the dimensions it contains or any information to suggest that it incorporates preferences.

The CHIP-CE began its development in 1995 and was designed to have the same domains as the CHIP-AE (Riley 2004a) so that health can be measured consistently throughout childhood and adolescence. (Riley 2004b) It is designed for children 6-11 years and there are parent and self complete versions. The parent version has an optional disorders domain in addition to the domains in the self complete version. There is some evidence of performance in a paediatric population, namely for test-retest reliability, internal consistency and practicality. It takes between 10 and 45 minutes to complete depending on the method and version used. Whilst there was evidence of a low rate of missing data, response rates were varied. Health status is reported at domain level and there is no incorporation of preferences into the scoring system.

The KIDSCREEN is a self or proxy complete health related quality of life questionnaire covering a wide age range (8-18 years). There are three versions (with 52, 27 and 10 items) which can be either proxy or self complete. Depending on the version used, it takes a relatively short time to complete (5-20 minutes). Scoring is by dimension and by global score for the short 10 item version. There is some evidence of psychometric performance in an adolescent population (internal consistency and test-retest reliability), but this is limited to one study (MAPI 2006). Preferences are not incorporated across the health states.

The TEDQL was developed to assess health related quality of life in children aged 3-8 years. It takes about 10-15 minutes to complete. There are 4 versions. There is limited evidence about its performance.

The YQoL-R was developed through a combination of interviews with adolescents, the literature and focus groups with adolescents, primary caregivers and child health and welfare professionals (Edwards 2002). There are two versions (the YQoL-S – 8 items and the YQoLR- 41 items) each with 4 domains. It is intended for 11-18 year olds and to be self complete. It adopts a very broad definition of quality of life and the authors argue that it should be youths themselves that define the concepts and items that are included in the measure which is why they took this approach to its development. There is some evidence of internal consistency and the time to complete is 2minutes for the YQoL-S and 15 minutes for the YQoL-R. This evidence is limited to one study. There is no incorporation of preferences across the health state descriptions and scoring is by summing all responses within domains (each scored on an 11 point scale) and then rescaling onto a 0-100 scale.

The HSCH-PS is still in the early stages of development and evidence of performance (test-retest reliability, inter-rater reliability, construct and convergent validity) is limited to one study (Saigal 2005). It has been developed for 2½ -5 year olds and has 12 domains, with 3 – 5 levels on each. It is designed to be completed by proxy, with a separate version for clinicians and parents. The scoring is by dimension, although the authors state they intend to do further work in the future to develop preference weights for the health states, making it suitable for use in economic evaluation. It has been developed from the dimensions of the HUI2/3 descriptive systems and the literature and the language has been adapted to make it suitable for pre school children.

In addition to the instruments listed in Table 3.1, Williams et al (Williams 2000) report the use of the CHQ Infant Toddler questionnaire for children aged 2 months to 5 years, but give no reference to its development and provide no evidence of its performance.

Matza et al (2005) demonstrated some evidence of construct validity in using the EQ-5D by proxy in a paediatric population, but it was extremely limited and poor evidence, only containing evidence for construct validity therefore it was not added to the table, especially as the EQ-5D is a measure designed for an adult population. The HUI3 was also primarily designed for adults, yet there was much more substantial evidence of its use in a paediatric population, especially as it is often used together with the HUI2 therefore this was still included.

Three measures were found that were developed for adults, but may be applicable to children, The Sickness Impact Profile (SIP) (Health Services and Sciences Research Resources Database), The SF-36 v2 and The SF-12. Whilst these measures claim to be suitable for adolescents, no evidence for use in a population under 16 years of age was found.

3.3.4 Purpose of instruments (Question 1b)

The majority of instruments have been developed to assess health related quality of life or quality of life with a view to this being useful information for evaluating the impact of diseases and treatments. The instruments vary in terms of the domains they include and some take a very broad view in terms of their definition of quality of life used. It could be argued that some of the domains in the instruments are less relevant for the purposes of health care decision making, as they are concerned with non health related factors, for example home environment, parental behaviour and family cohesion.
The exception to this is the HUI systems and the QWB which have been explicitly developed for the purposes of informing health care resource allocation decisions through economic evaluation, by providing preference information on the health states in the descriptive system, which, when combined with time in a health state, allows the calculation of quality adjusted life years (QALYs).

3.3.5 Evidence of performance in a paediatric population (Question 1c)

The majority of additional psychometric evidence found was for the Child Health Questionnaire (CHQ) and the PEDSQoL where there have been many studies evaluating the performance of these instruments in different populations, including cerebral palsy, obesity and cancer (CHQ) and rheumatology, asthma and diabetes (PEDSQoL). The evidence of the performance of the CHQ was good, with reasonable practicality (response rates from 50-70%, missing responses ranging from 0-18% per item and mean time to complete from 8 minutes (parent form) to 20 minutes (child form)) and good evidence of internal consistency (although in a couple of studies, this varied) and some evidence of reasonable test retest reliability, however the ICC (intra class correlation) ranged from 0.05 to 0.84 across studies.

The evidence of the performance of the PEDSQoL showed good practicality (low missing data), evidence of good internal consistency, however the evidence of inter rater reliability was fair, with agreement ranging from 0.358 to 0.635.

Evidence of practicality for the CHIPCE was mixed as missing data was low, time to complete varied from 10 to 45 minutes and response rates varied from 24-78%. There was some evidence of good internal consistency and some evidence of test retest reliability (ranging from 0.63 to 0.85 across studies).

The KIDSCREEN showed some evidence of practicality (time to complete ranging from 5 minutes (for the short version) to 20 minutes for the long version and some evidence of good internal consistency and evidence of test retest reliability, although estimates varied from 0.55 to 0.78.

3.3.6 Suitability for use as a PBM (question 1d)

None of the new instruments identified have any preference information in their scoring systems at present, which means they are unsuitable for use in economic evaluation in their current form. With the exception of the HSCS-PS, none have been designed with the purpose of being used in economic evaluation. The HSCS-PS was explicitly developed to be in line with the multi-attribute utility function approach to obtaining utilities. Although the preference weights have not yet been developed for this instrument, further work is planned to do this in the future. This instrument is designed for a very limited age range (2.5 -5 years) and although the developers state that there is potential for continuity across age ranges, with the use of the HUI2/HUI3 systems, this has yet to be tested. One of the main issues that will have to be addressed when obtaining values for the health states, is that this descriptive system contains 12 dimensions, which is above the usual capacity for processing of information (Miller 1956).

3.3.7 Adapting existing measures to become preference based (Question 1d)

With the exception of the HUI2, the instruments identified in this review are not suitable for adapting to become generic preference based measures of HRQoL. This is because the descriptive systems are either too large for valuation, or the content of the items includes non health related items, or the instruments have been developed specifically for use only with chronic populations and the structure of existing instruments is not ideal for valuation work. For example, the CHQ has 28 items in its shortest form and 98 in the long parent form, the CHIP-AE has 107 items plus 46 optional items, the KINDL has 40 items, the Nordic Quality of Life Questionnaire for Children has 75 items and the PEDSQoL has 45 items. These are all far too large to be amenable to preference based work. The Dartmouth Picture and Word COOP charts contains non health concepts, such as social support and family communication, the Generic Health Questionnaire contains relationships with parents and the Warwick Child Health and Morbidity Profile contains immunization. The Functional Status II and the Generic Health Questionnaire were both designed for use in chronic populations and the descriptive system may not be applicable to the whole paediatric population. To be amenable to valuation, instruments should ideally have a structure where there are between 5 and 9 dimensions and each has ordinal levels within it. With the exception of the HUI2, the instruments in this review do not have this structure and work would need to be done to adapt them. This is far from ideal for development of a PBM as reduction of items or dimensions could affect the breadth of coverage of a measure and important items or dimensions may be missed.

3.3.8 Evidence on what age it is feasible and reliable to elicit information about health related quality of life from children for the purposes of constructing a generic descriptive system. (Question 2)

Riley et al (Riley 2004c) report evidence that children can self report on their health, even as young as age 6, with more reliability and validity as age increases, as long as the questionnaire is age appropriate. Whilst this does not specifically give evidence that they can give information for the purposes of constructing a questionnaire, it does provide useful evidence that children are able to self report on their health and understand what is meant by health.

Although not used to develop a generic measure, Ronen et al (2001) report on the development of a measure for epilepsy where the objective was to develop a child-centred qualitative research methodology. 29 children were interviewed and stratified by age into 2 groups (6-9 years and 10-12 years). Group sizes were 3-5. Although this research was specifically about epilepsy and involved children with epilepsy, it does demonstrate that children are able to participate in research of this kind and provide information on their health related quality of life. The paper also states that they could not identify any work previously which has examined the methodologies to involve children in identifying items and domains relevant to their experience.

Graham et al (1997) developed the Child Quality of Life Questionnaire by interviewing 30 parents of children aged 10-14 years with chronic physical disorders, psychiatric disorders and mental retardation and 30 children with chronic physical disorders and psychiatric disorders. It is not stated what age the children were for the child interviews, although it could be assumed that they are approximately between 9 and 15 years old as this is who the measure was designed for and subsequently tested on. No details of the characteristics of the children are reported or how successful the interviews were and how they differed to those of adults. It is unclear whether the interviews were carried out separately from those with parents, whether the parents and children were related and whether the interviews were done on a 1 to 1 basis or in groups. Despite this lack of evidence about whether the children gave reliable information about health related quality of life, the authors subsequently developed a measure successfully so this could be taken as limited evidence that it is feasible to undertake interviews with children for this purpose.

Upton et al (Upton 2005) undertook some work to translate the PedsQl to a UK English version. As part of this process, they undertook cognitive interviews with 22 children to obtain feedback about how the items were understood by children. Modifications were made to the instrument in light of this. The age of the children interviewed was not reported. Although this is not strictly a paper describing how children have been used to develop an instrument, it does provide some evidence that they are able to participate in refining and developing the wording of health related quality of life questionnaires.

In the development of the Exeter Health Related Quality of Life Measure, Eiser et al (1999) undertook several interviews with children (age 6-13 years) as well as a review of the literature in order to develop the items for the questionnaire. No details are given about how many children were interviewed and whether they had any existing illnesses or conditions, or how much influence they had in the development of items.

In the development of the Quality of Life Profile – Adolescent Version, Raphael et al (1996) used adolescents from grades 9 -13 in Canada in groups of 6-8 to develop items for the measure. The adolescents were asked 'what does the term quality of life means to you?' and 'what are some areas of concern to adolescents?' Their responses were used to develop items by taking the most frequently raised items. The literature was also used in the development of the measure and the draft instrument was also taken back to the adolescents for validation work.

In the development of the Perceived Illness Experience measure, Eiser et al (1995) undertook semi structured interviews with 15 children and adolescents undergoing or having recently completed treatment for cancer. Patients were asked to recall experiences at points throughout their experiences, including diagnosis and return to school. From these, items were developed and underwent rating and further testing.

The PEDSQoL came from a measure that was originally developed for paediatric cancer but since then, a generic core scale has been developed from it and it has been used as a generic instrument. Children were used in the development of the cancer measure as the developers used the literature and interviews with parents and patients to generate an initial set of items (although no details are given on how many children were involved) but they were not involved in the development of the generic PEDSQoL, which was based on existing data (Varni 1999).

In developing the KINDL, Ravens-Sieberer (1998) et al developed a conceptual model of quality of life and then used these components in interviews with children. No details are given of the numbers or age of the children, however the paper reports that several school classes were used. Although children were involved in the development of the descriptive system, they were involved at a stage when items had already been developed for them to consider.

In developing the generic health questionnaire, Collier et al (1997) involved 80 children in schools and asked them to identify what made their lives good or bad. Their responses formed draft questions which were then piloted and refined further. The children were aged 6, 11 and 13 years.

The TedQoL is a self report measure for children 3-8 years old. The items in the instrument were developed from a review of the literature and previous experience of interviewing children, taking account of comments they had made about their interests and dislikes and their relationships with the people around them. The items were refined and added to as the development of the measure continued. The resulting measure was then tested for its psychometric properties. The TedQoL study also highlights the difficulty of obtaining self report quality of life data from children under 5 (Lawford 2001).

In the development of the YQoL-R, Edwards et al (2002) undertook in depth interviews with 33 adolescents aged between 12 and 18 years with and without disabilities. The adolescents were asked what was most important to them in their lives. They were sampled purposively based on gender, age, ethnicity, socioeconomics, sexual preference, disability, drop out and whether homeless and recruited until saturation was reached. Focus groups were then undertaken with adolescents and it was found that no new themes emerged that had not already arisen from the in depth interviews. Adult focus groups were also undertaken and different themes were emphasized to those of the adolescents. The study concluded that it is possible for adolescents to articulate their own views on the quality of their lives. It also demonstrates that the use of in depth 1 to 1 interviews was satisfactory in saturating the data as focus groups with the same age population revealed no new data.

Therefore there is some evidence that children can think about and describe their health and their quality of life, however this has often been done with the assistance of adults present or prompting from the literature. Some studies asked children to identify what was good and bad in their lives and children were able to do this, however this is very different from reporting on how health affects their quality of life. There is limited evidence to show that it is possible to ask children as young as 6 years old to describe their health and how it affects them and there is more evidence that this is possible, the older the children are.

3.3.9 Evidence on whether there are any age groups which share a stable HRQoL framework (question 3)

No evidence was found that looked at the issue of stable HRQoL frameworks across age groups or provided any data on this. The only potentially relevant papers were the development of the 16D (Apajasalo 1996a) and the 17D

(Apajasalo 1996b), which both made reference to Piaget (1969) which is concerned with the development of a child, rather than whether they share any stable health related quality of life frameworks. The 16D recognised 12-15 year olds as sufficiently homogenous in cognitive abilities and roles and the 17D 7-11 year olds based on this developmental literature.

Stein et al (1990) did some empirical testing of items in the FS II questionnaire to see what age range they could be applied to. They found that reliability was consistent for the following age groups: 5 years+, 4-8 years and 9years+. However, this is a very small piece of evidence and is not really the demonstrating that these age groups share the same frame work.

There was therefore no evidence on this issue and this is something that will require empirical testing.

3.4 Discussion

Since the review by McCabe, several new instruments were identified, adding 7 new instruments to the previous 20. None of the instruments was a PBM for children and the new measures all have quite high numbers of items, range in age of application from 3 to 18 years and differ in their coverage of domains. Some take a more general approach to quality of life and include items on relationships and the environment (YQoL-R/YQoL-S) whereas others such as the HSCH-PS focus on functioning. Some are in the early stages of development and there was limited evidence on their performance with the exception of the CHIP-CE which mainly had evidence of practicality.

None of the new instruments identified had any preference information in their scoring systems meaning they are unsuitable for use in economic evaluation in their current form. It would also be very difficult to adapt these existing instruments into a PBM as they often contain many items and are too large to be valued. In addition, they were not designed to be PBMs and hence do not have an appropriate structure for valuation. This is something that could be adapted but it is an inferior solution to developing a new measure suitable for its intended purpose, as adapting can significantly alter a descriptive system. The existing non PBMs were also not designed to be used in economic evaluation, therefore the content of the descriptive systems has been determined by other factors. They often contain items unrelated to health, such as income or family relationships which may not be of relevance to what a decision maker is interested in.

Overall, children have been used in the development of HRQoL measures but it is often at a stage when items have already been suggested and children are simply involved in refining the wording or testing out the items. In some cases, such as the Perceived Illness Experience and PedsQoL (original version), children with specific illnesses (cancer) have been used to develop the descriptive systems. This is different from developing a descriptive system based on a general paediatric population as those with a specific illness have been asked to think about this illness and how it affects their quality of life.

The majority of instruments listed in Table 3.1 have been developed using a combination of the literature and expert panels (containing paediatricians, psychologists, and parents). Even when children have been involved, it is rare that they are not prompted with pre existing items in the interviewing.
The only instruments where this is not the case and the children did not have a specific condition or illness they were asked to think about, are the YQoL-R, The Generic Health Questionnaire and the Child Quality of Life Questionnaire.

There was barely any evidence on the issue of stable health related quality of life frameworks, the only real reference to this issue was based on the 16D and 17D making reference to the developmental literature when deciding on appropriate age ranges for their questionnaires.

3.5 Conclusion

There is a range of generic paediatric quality of life measures available which have been developed for different purposes and adopt different definitions of quality of life. Most have some evidence of performance in a paediatric population, with the CHQ and the PEDSQoL having substantially more evidence than other measures.

No measures have been found that are suitable for adaptation and the HUI2 is the only existing PBM measure. The HUI2 was developed from existing literature and whilst children were involved in some of the development, this was in combination with their parents and they were asked to rate items that already existed, hence this is a top down approach to instrument development. For this reason, the HUI2 is less child focused. The HUI2 is also very functioning based and does not examine the consequences of something on quality of life, for example it contains items on hearing, speech and walking but does not contain items on the impact of these health problems on a child's quality of life.

Children have been involved in the development of some of the measures, but tended to be involved at a later stage, for example, when testing out the items. There is evidence that children can provide information about their health related quality of life, as low as 6 years of age, however there is little evidence on whether there are any stable health related quality of life frameworks between the paediatric age groups. If there were, this would give information about how to use appropriate age ranges when developing questionnaires. No measure has been found that has been explicitly developed for use in economic evaluation which uses children to develop the dimensions of HRQoL that it contains, hence there is a large gap in this area.

In general there is very limited guidance on key methodological issues facing the development of a new measure. The development of a new preference based measure of HRQoL described in this thesis will address some of these issues. The measure will be explicitly designed for use in economic evaluation and will include work designed to investigate whether there is a common HRQoL framework across age.

Appendix 3A: Checklist for judging the merits of preference based measures of health (Brazier 1999)

Practicality

- How long does the instrument take to complete?
- What is the response rate to the instrument?
- What is the rate of completion?

Reliability

- What is the test-retest reliability?
- What are the implications for sample size?
- What is the inter-rater reliability?
- What is the reliability between places of measurement?

Validity

Description 1; Content Validity
- Does the instrument cover all dimensions of health of interest?
- Do the items appear sensitive enough?

Description 2; Face Validity
- Are the items relevant and appropriate for the population?

Description 3; Construct Validity
- Can the unscored classification of the instrument detect known or expected differences or changes in health?

Valuation

- Whose values have been used?
- Assumptions about preferences?
  (a) What is the assumed model of preferences?
  (b) What are the main assumptions of this model?
  (c) How well are the preferences of patients/general population/decision makers likely to conform to these assumptions
- Technique of valuation
  (a) Is it choice based?
  (b) Which choice based method has been used?
- Quality of data
    (a) Are the background characteristics of the respondents to the valuation survey representative of the population?
    (b) What was the degree of variation in the valuation survey?
    (c) Was there evidence of the respondents understanding of the

References

Apajasalo, M., Rautonen, J., Holmberg, C., Sinkkonen, J., Aalberg, V., Pihko, H., Siimes, M.A., Kaitila, A., Makela, A., Erkkila, K. & Sintonen H. 1996b. Quality of life in pre-adolescence: A 17-dimensional health-related measure (17D). Quality of Life Research 5. 532-538.

Apajasalo, M., Sintonen, H., Holmberg, C., Sinkkonen, J., Aalberg, V., Pihko, H., Siimes, M.A., Kaitila, A., Makela, A., Rantakari, K., Anttila, R. & Rautonen J. 1996a. Quality of life in early adolescence: A sixteen-dimensional health-related measure (16D). Quality of Life Research 5, 205-211.

Asmussen, L., Olson, L.M., Grant, E.N. & et al. 2000. Use of the child health questionniare in a sample of moderate and low-income inner-city children with asthma. American Journal of Respiratory Critical Care Medicine. vol. 162, 1215-1221.

Brazier, J.E., Deverill, M., Green, C., Harper, R. & Booth A. 1999. A review of the use of health status measures in economic evaluation. 3[9].

Bruil, J., Maes, S., Le Coq, E.M. & Boeke, J. 1997. The development of the How Are You? (HAY), a Quality of Life Questionnaire for Children with a Chronic Illness. Quality of Life Newsletter. vol. 18 Sep-Dec.

Cadman, D., Goldsmith, C., Torrance, G., Boyle, M. & Furlong, W. 1986. Development of a health status index for ontario children; final report to the Ontario Ministry of Health.

Chan, K.S., Mangione-Smith, R., Burwinkle, T.M., Rosen, M. & Varni, J.W. 2005. The PedsQL: reliability and validity of the short-form generic core scales and Asthma Module. Medical Care. vol. 43, no. 3, 256-265.

CHQ. http://www.healthact.com/chq/chq_faq.asp
Accessed 01/03/2006

Collier, J. & MacKinlay, D. 1997. Developing a generic child quality of life measure. Health Psychology Update. vol. 28, 12-16.

Connolly, M.A. & Johnson, J.A. 1999. Measuring quality of life in paediatric patients. Pharmacoeconomics. vol. 16, no. 6, 605-625.

CRD. The University of York. Centre for Reviews and Dissemination.
http://www.york.ac.uk/inst/crd/crddatabases.htm
Accessed 24/01/2006

Edwards, T.C., Huebner, C.E., Connell, F.A. & Patrick, D.L. 2002. Adolescent quality of life, Part I: conceptual and measurement model. Journal of Adolescence. 25, 275-286.

Eiser, C., Havermans, T., Craft, A. & Kernahan, J. 1995. Development of a measure to assess the perceived illness experience after treatment for cancer. Archives of Disease in Childhood. 72, 302-307.

Eiser, C., Kopel, S., Cool, P. & Grimer, R. 1999. The Perceived Illness Experience Scale (PIE): reliability and validity revisited. Child: Care, Health & Development. vol. 25, no. 3, 179-190.

Eiser, C., Cotter, I., Oades, P., Seamark, D. & Smith, R. 1999. Health-related quality-of-life measures for children. International Journal of Cancer no. S12 87-90.

Eiser, C. & Morse, R. 2001. Quality-of-life measures in chronic diseases of childhood. [5;4].

Eiser, C., Vance, Y.H., Horne, B., Glaser, A. & Galvin, H. 2003. The value of the PedsQLTM in assessing quality of life in survivors of childhood cancer. Child: Care, Health & Development. vol. 29, no. 2, 95-102.

Furlong, W., Feeny, D. & Torrance, G. 1998. Multiplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report, McMaster University, 98-11.

Georgalas, C., Tolley, N. & Kanagalingam, J. 2004. Measuring quality of life in children with adenotonsillar disease with the Child Health Questionnaire: a first U.K. study. Laryngoscope, vol. 114, no. 10, 1849-1855.

Graham, P., Stevenson, J. & Flynn, D. 1997. A new measure of health related quality of life for children: preliminary findings. Psychology Health. vol. 12, no. 5, 655-665.

Health Utilities Index. http://healthutilities.biz/
Accessed 03/03/2006

Health Services and Sciences Research Resources Database.
www.nlm.nih.gov/nichsr/hsrr_search
Accessed 23/01/2006

Kaplan, R.M. 1989. Health Outcome Models for Policy Analysis. Health Psychology. vol. 8, no. 6, 723-735.

Landgraf, J.M. & Abetz, L.N. 1996. Measuring Health Outcomes in Pediatric Populations in Quality of Life and pharmacoeconomics in clinical trials, second edn, Spilker, B., (Eds), Lippincott-Raven Publishers, Philadelphia.

Landgraf, J.M. 1999. Measuring Pediatric Outcomes in Applied Clinical Settings: an Update about the Child Health Questionnaire (CHQ). Quality of Life Newsletter, vol. 23, 5-6.

Lawford, J., Volavka, N. & Eiser, C. 2001. A generic measure of quality of life for children aged 3-8 years: Results of two preliminary studies. Pediatric Rehabilitation. vol. 4, no. 4, 197-207.

Le Coq, E.M., Colland, V.T., Boeke, A.J.P., Bezemer, D.P. & Van Eijk, J.T. 2000. Reproducibility, construct validity, and responsiveness of the 'How Are You?' (HAY), a self-report quality of life questionnaire for children with asthma. Journal of Asthma. vol. 37, no. 1, 43-58.

MAPI research trust. Patient-Reported Outcome and Quality of Life Instruments Database. www.proqolid.org
Accessed 24/01/2006

Matza, L.S., Secnik, K., Mannix, S. & Sallee, F.R. 2005. Parent-proxy EQ-5D ratings of children with attention-deficit hyperactivity disorder in the US and the UK. Pharmacoeconomics. vol. 23, no. 8, 777-790.

McCabe, C.J. 2003. Estimating preference weights for paediatric health state classification (HUI2) and a comparison of methods. Ph.D.Thesis., The University of Sheffield.

Miller, G.A. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review. vol. 63, 81-97.

Patrick, D.L., Edwards, T.C. & Topolski, T.D. 2002. Adolescent quality of life, Part II: initial validation of a new instrument. Journal of Adolescence. 25, 287-300.

Piaget, J. & Inhelder, B. 1969. The Psychology of the Child New York: Basic Books.

Raphael, D., Rukholm, E., Brown, I., Hill-Bailey, P. & Donato, E. 1996. The Quality of Life Profile - Adolescent Version: Background, Descriptions and Initial Validation. Journal of Adolescent Health. vol. 19, 366-375.

Ravens-Sieberer, U. & Bullinger, M. 1998. Assessing health-related quality of life in chronically ill children with the German KINDL: First psychometric and content analytical results. Quality of Life Research. vol. 7, no. 5, 399-407.

Riley, A.W., Forrest, C.B., Rebok, G.W., Starfield, B., Green, B.F., Robertson, J.A. & Friello, P. 2004a. The Child Report Form of the CHIP-Child Edition: reliability and validity. Medical Care. vol. 42, no. 3, 221-231.

Riley, A.W., Forrest, C.B., Starfield, B., Rebok, G.W., Robertson, J.A. & Green, B.F. 2004b. The Parent Report Form of the CHIP-Child Edition: reliability and validity. Medical Care. vol. 42, no. 3, 210-220.

Riley, A.W. 2004c. Evidence that school-age children can self-report on their health. Ambulatory Pediatrics. vol. 4, no. 4, 371-376.

Ronen, G.M., Rosenb, Law, M. & et al. 2001. Health-related quality of life in childhood disorders: a modified focus group technique to involve children. Quality of Life Research. vol. 10, 71-79.

Saigal, S., Rosenbaum, P., Stoskopf, B., Hoult, L., Furlong, W., Feeny, D. & Hagan, R. 2005. Development, reliability and validity of a new measure of overall health for pre-school children. Quality of Life Research, vol. 14, no. 1, 243-257.

Speechley, K.N., Maunsell, E., Desmeules, M., Schanzer, D., Landgraf, J.M., Feeny, D.H. & Barrera, M.E. 1999. Mutual concurrent validity of the child health questionnaire and the health utilities index: An exploratory analysis using survivors of childhood cancer. International Journal of Cancer no. 95-105.

Spencer, N.J. & Coe, C. 1996. The development and validation of a measure of parent-reported child health and morbidity: The Warwick Child Health and Morbidity Profile. Child: Care, Health & Development, vol. 6, no. 22, 367-379.

Starfield, B., Bergner, M., Ensminger, M., Riley, A., Ryan, S., Green, B., McGauhey, P., Skinner, A. & Kim, S. 1993. Adolescent health status measurement: development of the Child Health and Illness Profile. Pediatrics. vol. 91, no. 2, 430-435.

Stein, R.E.K & Jessop, D.J. 1990. Functional Status II(R) A Measure of Child Health Status. Medical Care. 28[11], 1041-1055.

Sung, L., Greenberg, M.L., Doyle, J.J., Young, N.L., Ingber, S., Rubenstein, J., Wong, J., Samanta, T., McLimont, M. & Feldman, B.M. 2003. Construct validation of the Health Utilities Index and the Child Health Questionnaire in children undergoing cancer chemotherapy. British Journal of Cancer. vol. 88, no. 8, 1185-1190.

Trudel, J.G., Rivard, M., Dobkin, P.L., Leclerc, J.M. & Robaey, P. 1998. Psychometric properties of the Health Utilities Index Mark 2 system in paediatric oncology patients. Quality of Life Research. vol. 7, no. 5, 421-432.

Upton, P., Eiser, C., Cheung, I., Hutchings, H.A., Jenney, M., Maddocks, A., Russell, I.T. & Williams, J.G. 2005. Measurement properties of the UK-English version of the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. Health & Quality of Life Outcomes. vol. 3, (22) 1-7.

Uzark, K., Jones, K., Burwinkle, T. M. & Varni, J.W. 2003. The Pediatric Quality of Life Inventory in children with heart disease. Progress in Pediatric Cardiology. vol. 18, no. 2, 141-148.

Varni, J.W., Seid, M. & Rode, C.A. 1999. The PedsQL: Measurement Model for the Pediatric Quality of Life Inventory. Medical Care. 37[2], 126-139.

Varni, J.W., Seid, M. & Kurtin, P.S. 2001. Reliabilty and validity of the Pediatric Quality of Life Inventory Version 4.0 Generic Score Scales in healthy and patient populations. Medical Care. vol. 39, 800-812.

Varni, J. W., Seid, M., Knight, T.S., Uzark, K., & Ilona, S. 2002a. The PedsQLTM 4.0 generic core scales: Sensitivity, responsiveness, and impact on clinical decision-making. Journal of Behavioral Medicine. Apr 2002. Vol 25 (2): 175-193.

Varni, J. W., Seid, M., Smith, K.T., Burwinkle, T., Brown, J. & Szer, I.S. 2002b. The PedsQL in pediatric rheumatology: reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. Arthritis & Rheumatism. vol. 46, no. 3, 714-725.

Varni, J. W., Burwinkle, T.M., Jacobs, J.R., Gottschalk, M., Kaufman, F. & Jones, K.L. 2003. The PedsQL in type 1 and type 2 diabetes: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales and type 1 Diabetes Module. Diabetes Care. vol. 26, no. 3, 631-637.

Varni, J. W., Burwinkle, T.M., Rapoff, M.A., Kamps, J.L. & Olson, N. 2004. The PedsQL in pediatric asthma: reliability and validity of the Pediatric Quality of Life Inventory generic core scales and asthma module. Journal of Behavioral Medicine. vol. 27, no. 3, 297-318.

Vogels, T., Verrips, G.H., Verloove-Vanhorick, S.P., Fekkes, M., Kamphuis, R.P, Koopman, H.M., Theunissen, N.C.M. & Wit, J.M. 1998. Measuring health-related quality of life in children: the development of the TACQOL parent form. Quality of Life Research. vol. 7, 457-465.

Wake, M., Salmon, L. & Reddihough, D. 2003. Health status of Australian children with mild to severe cerebral palsy: cross-sectional survey using the Child Health Questionnaire Developmental Medicine and Child Neurology. vol. 45, no. 3, 194-199.

Wasson, J.H., Keirys, S.W., Nelson, E.C. & et al 1994. A Short survey for assessing helath and social problems of adolescents. The Journal of Family Practice. vol. 38, no. 5, 489-494.

Waters, E., Wright, M., Wake, M., Landgraf, J. & Salmon, L. 1999. Measuring the health and well-being of children and adolescents: A preliminary comparative evaluation of the Child Health Questionnaire in Australia. Ambulatory Child Health. vol. 5, no. 2, 131-141.

Waters, E., Salmon, L. & Wake, M. 2000a. The parent-form Child Health Questionnaire in Australia: Comparison of reliability, validity, structure, and norms. Journal of Pediatric Psychology. vol. 25, no. 6, 381-391.

Waters, E., Salmon, L., Wake, M., Hesketh, K. & Wright, M. 2000b. The Child Health Questionnaire in Australia: Reliability, validity and population means. Australian & New Zealand Journal of Public Health. vol. 24, no. 2, 207-210.

Waters, E.B., Salmon, L.A. & Wake, M. et al. 2001. The health and well-being of adolescents: a school-based population study of the self-report child health questionnaire. Journal of Adolescent Health. vol. 29, 140-149.

Waters, E., Stewart-Brown, S. & Fitzpatrick, R. 2003. Agreement between adolescent self-report and parent reports of health and well-being: results of an epidemiological study. Child Care Health and Development. vol. 29, no. 6, 501-509.

Williams, D.L. Geljins, A.C., Moskowitz, A.J., Weinberg, A.D., Ng, J.H., Crawford, E., Hayes, C.J. & Quaegebeur, J.M. 2000. Surgery for Congenital Heart Disease; Hypoplastic left heart syndrome: valuing the survival. The Journal of Thoracic and Cardiovascular Surgery. vol. 119, no. 4 part 1, 720-731.

Williams, J., Wake, M., Hesketh, K., Maher, E. & Waters, E. 2005. Health-related quality of life of overweight and obese children. Jama-Journal of the American Medical Association. vol. 293, no. 1, 70-76.

Chapter 4

Key decisions in developing a descriptive system

4.1 As discussed in Chapter 3, existing generic paediatric HRQoL measures are limited in terms of their use in cost utility analysis. With the exception of the Health Utilities Index 2 (HUI2) (Torrance 1996), they are not preference based, meaning that the calculation of QALYs from them is not directly possible. Although children were involved at some stage in the development of the HUI2, no preference based measure (PBM) to date has been developed primarily with children. The HUI2 is also based on a 'within skin' definition of health, in that it is very functioning and symptom based.

There are three stages to the development of a new PBM for children. The first is to construct the descriptive system, the second is to test and refine this and the final stage is to obtain values for all the health states defined by the descriptive system.

The first stage of developing the descriptive system raises a range of key decisions which must be made and are considered in this chapter:

1. Who is (are) the relevant population(s) for developing the content of the descriptive system?
2. What methods should be used to develop the descriptive system?
3. What concept of health/quality of life should be used?
4. What age range should be covered and is the health related quality of life (HRQoL) framework similar across this range?
5. Are there any developmental issues with children to be considered?
6. What are the constraints imposed by developing a PBM?

1. Relevant populations(s) for developing the content of the descriptive system

Many quality of life (Qol) measures in the past have been developed using a top down approach, in that they use the literature and views of experts, (generally clinicians) to develop the descriptive system (McColl 2005). Over the last 20 years or so, there has been a move towards involving patients and lay people in the development of QoL measures (Guyatt 1989). By involving the population who the measure is intended for, a greater content validity and relevance is likely to be achieved. It also gives a greater insight into the language and terminology used by the target respondents, ensuring the vocabulary of the instrument is appropriate(McColl 2005). This is a strong argument for using children as the population to develop the descriptive system.

The content of existing paediatric measures is generated from a mixture of literature reviews, expert opinion and interviews with relevant populations, including parents, children, paediatricians or other experts (Eiser 2001). No paediatric PBM to date has been developed purely from interviews with children. The only PBM, the Health Utilities Index 2 (HUI2) involved children in rating items taken from the literature. Similarly, in most general paediatric HRQoL measures, children have only been involved at a later stage, testing the items or wording. It is rare that children are used to develop the descriptive system itself without some other influence, for example prompting with existing literature or items, or the input of parents and/or clinicians.

Frequently descriptive systems for condition specific measures have been developed from interviews with patients, their carers, doctors and experts. It is believed that these people are the most appropriate and most informed to provide information on the dimensions of quality of life that are affected and important (McColl 2005). It is also easier to ask patients about their specific condition. A paediatric measure could be seen as a type of condition specific measure, in that it is specific to children and children are the patients.

Potential relevant populations that could provide the content for this new measure are parents/guardians, paediatric health care professionals, the general population, health care decision makers and children.

Parents have often been asked as proxies about the HRQoL of children, but there is evidence that parents views are affected by their own health status, knowledge, experience and expectations (Petrou 2003). There is mixed evidence in the literature about whether parent and child reports of HRQoL agree. There may also be areas or contexts of the child's life that are unknown or less well known to parents such as school (Matza 2004). For this reason, parents may not be the most appropriate source.

In a similar manner to parents, paediatric health care professionals will have indirect views, although informed about how health affects children. However, it is possible that their views too are likely to be affected by their own health status, knowledge, experience and expectations.

It could be argued that the general population have informed and relevant views, as everyone was once a child, and many will be parents. There is also an argument that they are representative of society and this measure is being developed for resource allocation decisions in society and therefore they would be an appropriate population to consult.

The purpose of the instrument is for NHS decision making, therefore including the views of those who are involved in resource allocation decisions (including decisions for children) may be appropriate. They make decisions and will have informed views about what is appropriate to consider in this context.

Including the views of children could be seen as most relevant as they are the population the measure is intended for. Content validity would be increased (McColl 2005) and the language and terminology would be more appropriate than using the views of others.

Some people may argue that children do not have the best knowledge of their own health or how their health affects them and parents have a better and more accurate idea (Matza 2004). However Drotar (2004a) 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. In addition, Matza (2004) reports that opinions vary, but it is generally estimated children can report domains of their own HRQoL from the age of 4 years for the most concrete domains, such as pain. Riley (2004) looked at evidence that children can report their health status and found evidence that they could, with adequate understanding, reliability and validity, especially over the age of 7 years.

There is a growing recognition that children have their own unique views and a right to express them in matters affecting them (United Nations) (Article 12 of the United Nations Convention on the Rights of the Child - UN 2007) and children are now taken more seriously as providers of data, with researchers actively seeking their views. Children are the population for whom the new measure is intended and the involvement of patients and lay people has been encouraged in developing PRO (patient reported outcome) measures (Guyatt 1989).

Although all the populations outlined above could provide information on what they feel are the appropriate and relevant dimensions for children for decision making, I would argue that the most relevant and appropriate population is the children themselves as they will be the users of the measure. This would also meet one of the tests of content validity of the FDA guidelines (Food and Drug Administration) which is how the items have been derived. Therefore this measure is developed with users (children) rather than experts.

2. What methods should be used to develop the measure?

In general, most quality of life (QoL) and patient reported outcome (PRO) measures have been developed using a top down approach, in that they use the literature, reviewing existing instruments and health surveys and the views of experts (usually clinicians or researchers) (McColl 2005) (Fitzpatrick 1998).

The main existing generic PBM for adults have all used this top down approach in the development of their descriptive systems. 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, including Euroqol, by surveying lay concepts (Van Agt 2005). The Quality of Well Being (QWB) drew its items mainly from an existing US Health Interview Survey and Social Security Administration Survey (Kaplan 1988), the SF-36 (from which the SF-6D is derived) used data from existing instruments (Short Form 36) and the Assessment of Quality of Life (AQOL) was developed from a literature review from 1970 and interviews and focus groups with 24 clinicians (Hawthorne 1997).

The only paediatric generic PBM, the Health Utilities Index 2 (HUI2) (Torrance 1996) used a review of epidemiological surveys and reviewed the literature to generate a large pool of potential attributes. They then used a sample of 84 child and parent pairs of the same gender living in the same household to rate these items, reducing the attribute pool from 15 to 6. The populations were sampled from schools in Hamilton, Ontario, Canada and were included as long as they could read, understand and speak English and the child was in grade 7 or 8 at school. A random sample was drawn from a central list of all students in Hamilton and stratified for gender and grade. (Cadman 1986).

Whilst children were involved in the rating stage along with their parents, the investigators made an expert judgement as to what attributes were relevant to the purpose for which the instrument was being developed when forming the initial list of attributes (McCabe 2003).

A contrasting approach to the top down methods of the EQ-5D, QWB, SF-6D, AQol and HUI2 is a bottom up approach, which takes the views of patients, 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 (McColl 2005). 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 (McColl 2005). In addition it will not be led or biased by inclusion of items from the literature or existing instruments, which may have been developed for a different purpose. It will also not be influenced by preconceptions about what should be included, for example as in the development of the HUI2 as outlined above.

Due to the many advantages of bottom up methods outlined above, this approach will be used in the development of the measure in this thesis.

3. What concept of health/quality of life should be used?

There are two separate issues to be considered, one is the distinction between quality of life and health related quality of life and the other is the actual definition or concept used.

Whilst there is no complete agreement in the literature on the definition, QoL is often defined as multidimensional, covering physical, emotional and social domains (Eiser 2001) and indeed many existing adult and paediatric instruments cover these areas (Dolan 1997) (Brazier 2002) (Landgraf 2005). Drotar (2004b) makes the distinction between QoL and HRQoL, by noting that HRQoL refers to the impact of an illness, treatment or health services policy on QoL (Stein 2004). Similarly, Matza (2004) notes a definition from the adult health outcomes literature that HRQoL is 'an individual's subjective perception of the impact of health status, including disease and treatment, on physical, psychological and social functioning.' QoL is much broader and encompasses other areas of a (child's) life, such as income, relationships and environment.

When developing a new measure, it is important to be clear on the definition of QoL/ HRQoL used. However, there is no agreed definition(s) in the literature, nor is there an agreed definition of health. The definition of health by the World

Health Organisation (WHO) states that health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity. (World Health Organisation). This has influenced measures developed since then, but other people have their own definitions and interpretations.

HRQoL can be viewed in many ways, ranging from a very impairment/symptoms based approach, to an approach based more on social functioning. For example, in the HUI2/3 the intent of the system is to assess aspects of impairment and disability (Torrance 1992). This approach can also be termed as 'within the skin'. The way in which these impairments affect your participation is a matter of choice.

In contrast, many of the psychosocial measures of quality of life used in trials are more to do with functioning and as a result are more subjective. Different functioning can have a different impact (on social functioning) depending on the individual (their response to the situation and their preferences). There are not many measures for paediatrics in the middle ground of this spectrum. As there is no agreed definition, the main point is to be clear over what concept and or/ definition of QoL/HRQoL is being used.

In constructing a paediatric measure of child health, Drotar (2004b) notes that the purpose of a measure, its content in terms of the domains considered and how and why they have been selected should be made explicit. In addition, Matza et al (2004) note that

"When designing a paediatric HRQoL instrument, it is important to ensure items correspond to experiences, activities and contexts that are directly relevant to the age of the sample"

The purpose of the new measure being developed is to aid resource allocation decisions in paediatric health care by providing information on the benefits of interventions in terms of their impact on health related quality of life. The descriptive system should contain dimensions of children's quality of life that are affected by their health, rather than dimensions that are not relevant for health care resource allocation decisions, for example items on clothes or toys which are related to income rather than health. In addition, the objectives of the UK NHS are health related rather than social welfare related and so the focus should be on health related dimensions (NHS 2008).

Therefore the concept used in developing the measure in this thesis is health related quality of life rather than quality of life, which is taken to mean the impact of health on a child's quality of life. This is also in line with the WHO definition.

4. What age range should be covered and is the HRQoL framework similar across the range?

Existing paediatric measures vary in terms of the age they are intended for, some covering a very wide age range and some much narrower (Eiser 2001). As found in the literature review in chapter 3, there is very little evidence about whether there are common HRQoL frameworks across age groups, for example, the dimensions of HRQoL important to a 16 year old may be very different to that of a 9 year old. This has long been recognised in the literature (Petrou 2003). It cannot be assumed that children as a whole group (age 0-16) are the same. This is an issue that can be tested empirically and will be examined in this thesis.

If HRQoL is stable across age, then descriptive systems for different ages should contain very similar dimensions.

5. Developmental issues with children

A similar issue related to the age of children is their development. In the earlier years especially, it can be difficult to disentangle developmental issues from health related quality of life issues. Relevant dimensions are potentially not the same across all age groups of children and many dimensions will arise due to different stages in a child's development. Children are often developing parts of their functioning for the first time at different ages and their baseline shifts as this develops (Stein 2004). This is especially true for children under 5 years old when they are learning to crawl, walk, talk, feed themselves, dress themselves and care for themselves, for example using the toilet independently or washing. The developmental milestones can vary by age quite significantly, which makes it difficult for health status measurement. For example, the HUI2 contains a question about being able to walk unaided. This is not meaningful for children who cannot yet walk because they are too young, rather than not being able to walk on account of their health. Similarly there is a question on whether the child can communicate, which is not relevant if the child has not yet learnt to talk. From about the age of 5 years old most of these issues are no longer relevant as children can usually walk, talk, eat and dress themselves.

6. What are the constraints imposed by developing a PBM?

This thesis is concerned with developing a PBM for use in economic evaluation and this imposes an additional consideration to conventional non PBMs. The main constraint is that the health states defined by the system should be amenable to valuation. This means there is a need for a health state classification system and that there is a limit to the number of dimensions/levels that it can contain. Descriptive systems that have dimensions with ordinal levels fit these criteria well.

The most widely used generic descriptive systems range from five to nine dimensions (Brazier 2007). 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 first part of this research concentrates on identifying all the relevant dimensions of HRQoL for children and then subsequently reducing these to meet valuation constraints.

4.2 Approach taken to the development of the new measure

The aim was to develop a descriptive system that is relevant for its purpose (paediatric health care resource allocation decisions) and relevant to the population it will be used on (children). Therefore it should contain dimensions of HRQoL that are relevant to children. As it is a generic measure, it should include the breadth of ways in which health affects children's lives across all conditions and severities. Most of the issues discussed previously can be related to the purpose of the measure, which should drive the decisions made concerning its development. It should contain dimensions related to the impact of children's health problems, which means that children are ideally placed to determine what these are and also the breadth of these dimensions. In addition, taking a bottom up approach ensures that the measure is developed according to its purpose, rather than incorporating information from previous research that may have had different aims and objectives.

To achieve a greater content validity and relevance to children and also to ensure appropriate language and terminology, a decision was made to use children as the only population to develop the content of the descriptive system.

Whilst the purpose of this measure is to inform paediatric resource allocation decisions, it should be noted that in the UK (and also elsewhere), no single budget exists for child health care. Instead, decisions tend to be made across varying age groups. It has been argued that the content of a preference based paediatric measure (in terms of the dimensions included) should be determined, at least in part, by whether the measure will be used to inform resource allocation decision across or within age groups (Petrou 2003). For example, whilst the development of age specific measures due to the potential differences in relevant attributes at different stages of a child's life will mean the measure is appropriate for the population (as in this work), it could mean we would lose the ability to make resource allocation decisions across individuals of different ages.

An alternative approach to developing different measures for different age groups is to develop a single measure that can inform resource allocation decisions across all age groups (including adults). This would be something that would be very hard in practice, due to different dimensions being relevant or irrelevant at different stages of people's lives and also the terminology not being age specific (Petrou 2003).

Perhaps one solution to this problem may be that the common metric of measurement (the QALY) should be what to focus on. For example, bespoke descriptive systems relevant to the population could be developed and then valued using a common method (for example time trade off). In that way, all measures would be on the same scale but the descriptive systems would be valid for the population.

4.2.1 Sample

The new measure is intended to be generic therefore children with as wide a range of health problems as possible were included. There are two main areas where children could be sampled from; schools or hospitals.

Sampling children in mainstream schools would include those who have had or have chronic or acute conditions for example diabetes, asthma, eczema or flu. Those children who were hospitalized, too ill to come to school, and those with severe learning difficulties who attend special schools, i.e. the more severe end of the population would not be included.

Sampling children from a hospital would in contrast, get the more severely affected end of the paediatric population, but may miss a whole range of more minor conditions or conditions that rarely require hospital visits. In addition, interviewing children about their health and how it affects their lives is a very sensitive area and it is difficult ethically to interview children who have a terminal illness for example.

A decision was therefore made to sample children from schools as a greater breadth and range of health problems could be achieved, however it was accepted that children at the sicker end of the population would be missed. Further work testing the measure in a hospital setting could help to examine the extent of this limitation.

As there was only one researcher on this project, sampling could only be done within the Sheffield area, as it was infeasible to sample children from schools from other areas in the country due to time and resource constraints. Sheffield is a large city and careful sampling of schools ensured that the population was representative of the general paediatric population.

Children aged 7 to 11 years were chosen as this avoided some of the developmental issues that would occur with children under 5, kept the focus within a reasonably tight age span and also mirrored the school system in that primary or junior schools have children between these ages. The research also included work to test whether 2 age groups within this range share a similar HRQoL framework (7-9 years, 9-11 years).

4.2.2. Coverage and content

The measure developed in this thesis is for paediatric health care resource allocation decisions, therefore it needs to capture how health affects children's everyday lives by capturing changes that are important, appropriate and relevant to children. It is being developed as a HRQoL measure, in that the dimensions are related to health and not other areas such as income or family environment. It explores the consequences of a health problem on a child's life, rather than being symptom based. Therefore to be a dimension of HRQoL it must be a dimension that is affected as a result of health (and not of some other consequence, e.g. sibling relationship, income, or environment).

For adults, it is agreed that a patient should self complete a HRQoL measure as they usually have the best knowledge of their HRQoL at any particular point in time (Matza 2004). If this is not possible, for example in vulnerable groups such as the very elderly, people with severe mental health problems, learning disabilities and children, completion can be done by proxy, such as a relative, carer or parent. There is still a debate about whether children are able to self rate their health (Eiser 2001), however there is an increasing amount of evidence that children can provide reliable and valid self assessments about their physical symptoms and emotional well-being (Matza 2004) (Landgraf 2005). As this measure is being developed using bottom up methods, it is much more likely that children will be able to self complete as the language and terminology will all be determined by children. If a child can provide reliable and valid data, then self report is optimal (Matza 2004), therefore the new measure is being developed with the intention that children will self complete.

4.2.3 Issues in researching and working with children

There is evidence that children can provide information about their HRQoL. Riley et al (2004) report that children can self report on their health, even as young as age 6, with more reliability and validity as age increases, provided the questionnaire is age appropriate. Whilst this does not specifically provide evidence that they can give information for the purposes of developing the content of a descriptive system, it does suggest that children are able to report on their own health and may understand what is meant by health.

The location of where research takes place is likely to influence the way children respond. School based research is highly cost effective, as there are high participation and completion rates. Responses are also less likely to be influenced by parents. However, it is important that the children do not feel the research is a test and that they are being judged or that there are any right or wrong answers.

Children as a group are immensely diverse as there are a huge range of differences between children including age, ethnicity, gender, social background and characteristics, for example how shy or confident the child is, how good they are at communicating and whether they have any hearing or visual problems to name just a few. The researcher must be aware of these differences and be able to adapt appropriately if necessary.

There is a choice to be made between using focus groups or individual interviews with children. The advantage of focus groups is that children can feel more comfortable and feed off each others ideas, however if the material being discussed is sensitive, it may be better to do individual interviews. In addition, pre existing dynamics within the groups of children selected can have an influence. Individual interviews can feel uncomfortable for shy children and many children will feel nervous, however this is something that can usually be addressed by the researcher. Focus groups can also lead to inhibitions in raising issues that some children may feel are important, yet others do not and so children are unwilling to share them. Individual interviews give the opportunity for complete privacy and the opportunity for children to raise any issues that they wish.

Thomas et al (1995) found that there was no difference in terms of the depth of data generated between focus groups and interviews, however their study was conducted on an adult population therefore this may not be generalisable to children. 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 (McColl 2005). The practical issues of working with children and the potential sensitivity of the topic meant that individual interviews were more appropriate for this study.

It is vital to ensure that children are fully informed about the research and have given their informed consent before they participate. Care must be taken to create a friendly and relaxed environment, so that children do not feel intimidated or uneasy. It is also important for the researcher to be able to accommodate the range of children that may take part, for example concentration levels or cognitive abilities can vary hugely. In addition, some children may be very shy or nervous whereas others may find it difficult to focus on the task.

If children know what is going on and why, they will feel more comfortable going into an interview, which results in a much more successful interview. Therefore the preparation before the interview is key – children should be reassured that there are no right or wrong answers and that you are interested in hearing their views and that they are the expert. The child should understand that they can stop at any time and it is a useful idea to agree a signal with them, for example tell them what to say if they want to stop. They should always be given the opportunity to ask questions before, during and after the interview and understand that they can do so at any time. They must be reassured that the interview is confidential unless there are disclosure issues and this should be explained as part of the consent process. All these points should be provided either as an information leaflet, or verbally by the researcher, or both, prior to asking their consent to take part.

It is good practice to include a warm up exercise before undertaking the main research, as this can help put the child at ease, for example a short game or creative task. This gives the opportunity for the child to get to know the researcher and to relax and feel comfortable talking to them.

General principles for interviewing children are similar to those used when conducting any type of qualitative interview. In particular, short open questions should be used and closed questions avoided. For example closed questions that start with ‘do you’ or ‘did you’, invite short categorical responses. Also, a closed question may result in acquiescence bias by implying to the respondent that a positive response is required (McColl 2005), for example, "does it affect your schoolwork?". This is particularly an issue with children as there is a danger they will think you are looking for a specific answer. A better question would be "how does this affect your schoolwork". Open questions generally start with 'what' or 'how' and usually allow greater depth to be gained from the respondent.

It is important not to make assumptions about what the child is saying during an interview and not to react in a positive or negative manner to what they are saying. Instead, display signs of active listening which shows that you are interested in what they are saying but are not passing any judgement on anything.

Unfortunately, there is not much guidance in the literature about one to one depth interviewing with children for the purposes of research, however the general principles of qualitative interviewing and working with children which are outlined above were used to inform the interviews.

4.2.4. Study Design

Mixed methods were used to develop the measure in this thesis. The first stage involved qualitative methods which were appropriate for this type of research question, exploring the breadth of the ways in which health affects children's lives. The aim of the first stage of the research was to identify all the dimensions of HRQoL relevant to children and reduce them at a later stage to meet the size and design required for valuation.

There are several stages to qualitative research design (Ritchie 2005). Firstly you should review your research question and decide what type of qualitative data you require. In this case, 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 was appropriate because you need to generate the data by asking children about how their health affects them, this is not something that would be naturally occurring and you could observe, for example if you wanted to investigate child relationships, you could just observe children in a natural setting.

Secondly, you should choose your data unit and time span for research. In this case, this was case studies of individuals at a point in time (cross sectional). Thirdly, the data collection method for generated data should be decided. In this case it was individual interviews, which would be recorded and transcribed verbatim. Individual interviews were chosen because of the sensitivity of the topic and 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.

A semi structured interview was decided on, as it uses open ended questions and has a reasonably loose structure which is appropriate for this type of research. It is exploratory in nature, yet seeks to find out about specific issues (McColl 2005). These type of interviews generally use a topic guide which defines the general areas to be explored and aids the researcher by ensuring that each interview covers the same general questions.

A sampling strategy needs to be determined. Here, children aged 7-11 years old were sampled from schools with a range of socio economic characteristics. A sampling strategy of purposive sampling, based mainly on level of health was applied in order to include the breadth of experiences of health. In qualitative research, the aim is not to achieve statistical representativeness, but to 'identify specific groups of people who.... possess characteristics....relevant to the phenomenon being studied (Ritchie 2005). In this case, this is experience of health problems. Sample size is not pre determined and interviews are undertaken until saturation is reached, which is when no new themes emerge.

Finally, an analytical approach and data management tool need to be determined. Thematic content analysis was chosen as this is appropriate to the research question and the type of data collected. Thematic content analysis is an approach whereby themes are generated by the data and the content of the interviews is analysed. This approach was chosen as it suits the approach of generating dimensions from this type of interview data. This analytic approach was aided by the use of the software package NVIVO. (NVIVO)

In summary, the following key decisions were made which informed the design of the measure: Children were the relevant population and a bottom up approach would be taken using semi-structured individual interviews as they would be more appropriate given the nature of the topic and the practicalities of managing children. The concept used was HRQoL which is the impact that a health problem has on your life. The purpose of the interviews was to determine how their health affected their lives in order to develop relevant dimensions of HRQoL for this population. The age range to be covered was 7-11 years old to avoid any developmental issues and to keep the focus within a reasonably tight age span. Work was also included to test whether the age groups 7-9 and 9-11 years share similar frameworks. The constraints of a PBM influenced the design in that the health state classification system was developed with dimensions and levels and the interviews were designed to generate these dimensions.

References

Brazier, J.E., Roberts, J. & Deverill, M. 2002. The estimation of a preference based measure of health from the SF-36. Journal Health Economics. 21 (2), 271-292.

Brazier, J.E, Ratcliffe, J., Salomon, J. & Tsuchiya A. 2007. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford University Press.

Cadman D, Goldsmith C, Torrance G, Boyle M, Furlong W. 1986. Development of a Health Status Index for Ontario Children. Final Report to the Ontario Ministry of Health Research Grant. DM648 (00633).

Dolan, P. 1997. Modelling valuations for EuroQol Health States. Medical Care, 35, 1095-1108.

Drotar, D. 2004a. Measuring Child Health: Scientific Questions, Challenges, and Recommendations. Ambulatory Paediatrics. 4, 353-357.

Drotar, D. 2004b. 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.

Eiser, C. & Morse R. 2001. Quality-of-life measures in chronic diseases of childhood. Health Technology Assessment. 5(4).

Fitzpatrick, R., Davey, C., Buxton, M.J. & Jones, D.R. 1998. Evaluating patient-based outcome measures for use in clinical trials. Health Technology Assessment. 2(14).

FDA (Food and Drug Administration), U.S. Department of Health and Human Services. February 2006. Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Draft Guidance.

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.

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.

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.

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

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.

McCabe, C. May 2003. Estimating preference weights for paediatric health state classification (HUI2) and a comparison of methods. Ph.D.Thesis. The University of Sheffield.

McColl, E. 2005. Developing questionnaires. In Fayers, P. & Hays, R. (Ed.). Assessing Quality of Life in Clinical Trials. 2nd edition. Oxford University Press.

NVIVO7 QSR International

National Health Service
http://www.nhs.uk/Pages/homepage.aspx?WT.srch=1
Accessed 21/11/2008

Petrou, S. 2003. Methodological issues raised by preference-based approaches to measuring the health status of children. Health Economics. 12: 697-702.

Riley, A.W. 2004. Evidence that school-age children can self-report on their health. Ambulatory Pediatrics, 4(4), 371-376.

Ritchie, L., Lewis, J. (Eds.) 2005. Qualitative Research Practice; A Guide for Social Science Students and Researchers. SAGE publications, London.

Short Form 36. http://www.sf-36.org/tools/SF36.shtml#VERS2
Accessed 13/03/2008

Stein, R.E. K. 2004. Measurement of Children's Health. Ambulatory Pediatrics. 4, 365-370.

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.

Torrance, G.W., Zhang, Y., Feeny, D., Furlong, W. & Barr, R. 1992. Multi-Attribute Preference Functions for A Comprehensive Health Status Classification System. McMaster University, Centre for Health Economics and Policy Analysis Working Paper No. 92-18.

Torrance, G.W., Feeny, D.H., Furlong, W.J., Barr, R.D., Zhang, Y., & Wang Q. 1996. A multi-attribute utility function for a comprehensive health status classification system: Health Utilities Mark 2. Medical Care, 34(7), 702-722.

United Nations
http://www.article12.org/pdf/UNCRC%20Official%20Document.pdf
Accessed 08/02/07

Van Agt, H, Bonsel, G. In Kind P, Brooks R, Rabin R (eds). EQ-5D concepts and methods, a developmental history. Springer; 2005.

World Health Organization. 1948. Constitution of the World Health Organization. Basic Documents. WHO, Geneva.

Chapter 5

Developing the Dimensions

5.1 Introduction

This chapter describes the first stage in the development of the new HRQoL measure which was to develop the dimensions that would form the basis of the descriptive system. As justified and decided in chapter 4, only children aged between 7-11 years old are being used to develop the descriptive system and children were accessed via schools. The primary aim of this stage of the research was to develop a long list of dimensions relevant to this population that covers the breadth of ways in which health affects children's lives. A second aim was to explore whether there was a common HRQoL framework across age within this group by examining two age sub groups.

5.2 Methods

Prior to undertaking the main qualitative interview work, a pilot study was carried out to pilot all the processes and methods of the proposed qualitative research. Materials were drafted for the consent process (letter to parents, information leaflet for parents, information leaflet for children and consent form for parents) and a draft topic guide was developed (Appendix 5.A) which was used to undertake the pilot interviews.

5.2.1 Piloting

The aim was to pilot the whole interview process including consent materials, the topic guide used for the interviews, recording equipment and materials used during the interview.

5.2.2 Methods

Staff at ScHARR were contacted to ask if they had any children aged 7-11 years and if so, for their consent to interview their child. A date and time was arranged with those who responded and the consent form was signed by the parent. In addition, parents were provided with the information leaflet and letter to parents that were to be used in the main study in schools and asked for their comments on these as well as on the consent form. All interviews were undertaken at Regent Court at The University of Sheffield. The child was asked to read the information leaflet for children designed for the main study. It was explained to them that they were just testing the process and it would be eventually carried out in schools and they were invited to give feedback on what they thought of the leaflet. Once they had read and understood this and had the opportunity to ask questions, they were asked for their consent (verbally) to take part in the interview. The interview was undertaken on a 1 to 1 basis with the parent present in the room but not taking a role. Stickers and name badges were used as a warm up exercise before undertaking the interview. The draft topic guide was followed which was designed to first ask children about their health in general and any health problems they had and then to describe how these affected their lives. Prompts were used to gain greater depth and every health problem the child had was explored. Digital recording equipment was used to record the interviews and children were given a certificate for taking part at the end, in the child's choice of colour.

Ethical approval was obtained for this part of the work through the School of Health and Related Research (ScHARR) ethics committee.

5.2.3 Results

Four interviews were carried out. Children were aged 8, 8, 10 and 11 years. Three were female and one was male. All interviews were carried out successfully and all children appeared to enjoy taking part. Interviews lasted 20 -25 minutes. The recording equipment worked very well with no problems and all data was successfully uploaded to the computer. Children did not mind the recording equipment being used as long as it was explained why it was necessary beforehand. The warm up exercise worked very well and some children asked if the interviewer could wear a name badge in case they forgot their name. There was a tendency for children to want closed questions rather than open ones and they occasionally got a bit impatient with probing, as they felt they had already given an answer, even though more depth was obtained in this way.

There was often confusion between the concepts of health and healthy – i.e. some children thought their health was very good because they did not eat many sweets and drank lots of water.

5.2.4 Conclusions

The piloting work made it clear that it is important to make the distinction between health and healthy at the beginning of an interview, perhaps by giving an example of a condition, such as asthma, to make sure children are clear about this. This could then be followed by asking about any health problems they may have. Following this, children can be asked to think through their environment and how their health affects their lives. It was also important to try and avoid closed questions and continue with the probing, as more depth and information was obtained. Explaining to children at the beginning of the interview what the research was about and what they were being asked to think about was very important and made a big difference to how successful the interview was.

All materials for the interviews worked well (stickers and certificates). The warm up exercise to relax prior to the actual interview was proved essential to improving the quality of interview obtained as it was vital that children felt relaxed and comfortable.

5.3 Main research study

5.3.1 Access

The schools were chosen in collaboration with the Children and Young People's Directorate at Sheffield City Council specifically to represent the diversity of children in Sheffield in terms of ethnicity and social class. Their recommendation was to choose 2 schools that between them covered the two main types of state school found in Sheffield, namely affluent middle class areas and less affluent inner city schools. The two schools identified as the most representative were Firs Hill Community Primary School and Hunters Bar Junior School. The Head Teachers of both schools were contacted and both agreed to take part in the research. The characteristics of the schools are given in Tables 5.1 and 5.2 togther with comparative Sheffield and National data. Both schools have an equal mix of gender and sit either side of the median in terms of a key indicator of deprivation - eligibility for free school meals, and total special educational needs. The percentage of ethnic minority children and percentage with a first language other than English are above the median for Sheffield and the UK in both schools and therefore adequately cover the range of children in terms of age, gender, sex and ethnicity.

Table 5.1: Characteristics of the schools - gender

	Gender	Y3	Y4	Y5	Y6	Total
Firs Hill+	% boys	60.66%	47.46%	56.67%	53.33%	54.58%
Hunters Bar+	% boys	42.39%	48.91%	47.83%	46.24%	46.34%
LEA+	% boys	51.32%	51.22%	50.78%	45.76%	51.13%
National++	% boys					51.22%

Table 5.2: Characteristics of the schools - general

	Eligible for Free School Meals %	Total SEN* %	Ethnic Minority (NCY** 1 to 6) %	First Language other than English (NCY** 1-6) %
Firs Hill Community Primary+	25.00%	27.00%	83.10%	68.90%
Hunters Bar Junior+	9.50%	16.00%	36.00%	18.20%

LEA^{+}	19.1%	21.0%	21.4%	13.5%
Median for Sheffield^{+}	13.64%	21.18%	11.25%	2.33%
National^{++}	16.98%	17.28%	20.6%	12.5%

$^{+}$Source: Pupil Level Annual School Census (PLASC) as of January 2006. PLASC is a census of all children in the school system. Data provided by The Children and Young People's Directorate at Sheffield City Council

$^{++}$ Source: http://www.dfes.gov.uk/rsgateway/DB/SFR/s000682/index.shtml
Accessed 20/12/2007
http://www.dfes.gov.uk/rsgateway/DB/SFR/s000744/UPDATEDSFR30_2007.pdf
Accessed 20/12/2007

*SEN (special educational needs)

**NCY (National Curriculum Year)

A criminal records bureau (CRB) disclosure check at the enhanced level was completed in January 2006 for the researcher (KS) carrying out the interviews and the Children and Young People's Directorate at Sheffield City Council reviewed the proposed research to check the ethical considerations.

5.3.2 Age

To test out whether children within the 7-11 years age group share a common health related quality of life framework, each school was split into two groups; Y3+Y4 (7-9 years old) and Y5+Y6 (9-11 years old). This split was chosen as creating 4 groups (one for each year) would be too difficult to manage.

5.3.3 Sampling

All parents of children in both schools were sent a letter explaining the nature of the study, an information leaflet and a consent form asking for their consent for the researcher to approach their child to take part as well as completing information on their child's gender and ethnicity. (Appendix 5.B, 5.C and 5.D) Parents were asked to rate the health of their child using five levels (excellent, very good, good, fair and poor) and then return the form to their child's school if they were willing to consent. The levels were taken from Q16 of the HUI2 which is not actually part of the HUI2 descriptive system but is a general question. From the returned consent forms children were sampled purposively.

The primary sampling criteria was age, followed by level of health (as rated by the parent) to ensure that the views of a full range of health were covered. Level of health was used as a sampling criteria, as this was expected to have a large impact on the views children had about their HRQoL. The question used to determine this was taken from Q16 of the HUI2 as there was evidence of good inter rater reliability between parents and children with this question in a paediatric population in the UK (own analysis of a previous study, not yet published). Secondary criteria were gender and ethnicity. The two groups described above were sampled and analysed independently to explore whether they shared the same HRQoL frameworks, in terms of the dimensions they generated. The sample was purposive in order to ensure that the full breadth of views was included.

5.3.4 Interviews

Children were interviewed in school in a suitable place such as the library or the dining room, so they could concentrate and no one else could hear them. Children received an information leaflet and read through this with the interviewer and had the opportunity to ask questions. A warm up exercise was undertaken to make the child feel relaxed and comfortable which involved making a name badge with a sticker. Whilst it was decided that 1 to 1 interviews would be preferred (as discussed in chapter 4), it was felt important to offer children a choice and so children were asked if they preferred to do the interviews in a one to one situation or a small group, as talking about health problems can be a sensitive area.

A semi structured interview was held, which asked about any health problems children had and how they affected their life, using a topic guide (Appendix 5.E) developed to include probing to make sure both acute and chronic health problems were recognised as valid and that all areas of children's lives were considered, for example, how their health affected their lives both at home and school. The format of the interview was firstly to ask the child about any health problems they had and then ask for some basic descriptive information, in order to engage them and get them thinking about their health. Then, when the child had described the nature of the health problem, further questions were asked about how their health affected their lives. Probing was necessary in the interviews, both for making sure the child thought about how their health affected them in all areas of their lives such as home and school, and also to make sure all their health problems were included. This was achieved using prompts such as 'how does this affect you at school', 'have you been to the doctor', 'do you take any medicines' and 'have you been off school recently due to your health'. The interviews were designed to understand and explore the consequences of a health problem on the lives of children, rather than just the symptoms a health problem may produce. These consequences then form the dimensions of "HRQoL". The use of closed questions was avoided, even though children prefer this type of questioning, since it is poor interview technique and limits the data obtained. Instead, open questions, for example "how did that affect you" and "why did it make you feel like that" were used, in order to encourage children to give more depth and explanation to their answers.

All interviews were carried out by the author, recorded and transcribed verbatim and continued until saturation was reached. The position of the researcher was neutral, in that no influences from existing paediatric measures of HRQoL were brought to the interviews. It was left to the children to determine how they felt their health affected their lives. Children understood that they were free to stop at any time if they wished and return to normal lessons.

5.3.5 Analysis

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. Thematic content analysis (analysis of the content of interviews and looking for themes) was undertaken using Framework, an approach developed by the National Centre for Social Research (Ritchie 2005). This provides a systematic thematic way of summarizing and classifying data. It involves several stages, the first is to review the data by rereading and relistening to all the interviews in order to be become more familiar with the data. During this process, recurring themes and ideas were identified as they occurred and a thematic framework was devised by grouping these into main themes and sub themes. All interviews were then coded according to this framework, with the aid of NVIVO software (NVIVO). The data were charted, producing a matrix of sub themes and respondents, where each row in the matrix was a respondent and each column was a sub theme. This matrix summarized and synthesized each sub theme, taking care to retain the terminology and language of the children. All data were charted in this way. Each sub theme was then reviewed and explanations behind the affected areas of HRQoL explored. Relationships between themes and sub themes were investigated and mapped and dimensions were generated by exploring the consequences of a health problem on a child's life. For example, an ear infection may mean that the child cannot hear well, which then means that they cannot hear as well in lessons, affecting their schoolwork. In this case, school work would become a dimension. Another example would be a child who has a problem with their leg and consequently can't play football. This leads them to feel angry but also not being able to join in the activities they would normally do. In this case, there would be 2 dimensions; angry and joining in activities. The aim of the analysis was to look for the end consequence of a health problem and not just the symptoms it might lead too. The interviewing was undertaken in such as way as to probe children for these consequences and get the depth and reasoning behind them. It is these end consequences which become the dimensions of health related quality of life. There could be many difference reasons as to why the dimensions occurred, however it was the end consequence (or the dimension) that was of interest, not what caused it. Dimensions were intended to be mutually exclusive.

The analysis was overseen by an experienced and independent qualitative researcher, (Dr Alicia O'Cathain, ScHARR, The University of Sheffield) who reviewed the charting and mapping process. Further details of the Framework process are given in Appendix 5.F.

5.4 Results

5.4.1 Sampling

There were 232 children in Y3-6 inclusive in Firs Hill school and all were given a letter, information sheet and consent form (distributed by their class teacher to take home) at the end of April 2006. Approximately 2 weeks later, 22 consent forms had been returned to the school (9.5%). All were distributed amongst the top three categories of health and 1 in the fourth category. Therefore targeting was done by the Head Teacher to encourage those parents of children with known health problems to take part (to try and sample from those in the fourth and fifth categories.) 1 child was recruited in this way, with level 5 health.

There were 369 children in Y3-6 inclusive Hunters Bar School and all were given a letter, information sheet and consent form (distributed by their class teacher to take home) in May 2006 (approx 19th). 129 forms were collected on 8th June (35%) and 3 more were collected on 14th June. 131 in total consented (some had problems with missing signatures and went back to parents and 1 was not returned). After sampling based on level of health and trying to balance for gender and ethnicity, the head teacher was asked to identify any children whose parents had consented from the remaining pool of non interviewed children that they thought would be useful to interview as they had known health problems. This led to 6 more interviews (4 in Y3/4 and 2 in Y5/6)

In total, 17 interviews were carried out on Y3/4 Firs Hill children, 24 on Y3/4 Hunters Bar children, 5 on Y5/6 Firs Hill children and 28 on Y5/6 Hunters Bar children. This was 74 interviews in total on 75 children (1 interview was carried out as a pair). Tables 5.3 to 5.6 below show the sampling grids of health levels (1-5) and gender/ethnicity split for each school and age group.

Table 5.3: Sampling Grid Y3+Y4 Firs Hill n=17

Health Level	M1	M2	M3	M4	M5	MO	F1	F2	F3	F4	F5	FO
1			1				1		2			
2	1		2				1		1			
3				1		1	1		3			
4			1									
5									1			

(M=male, F=female, 1=white, 2=mixed/dual heritage, 3=asian or asian british, 4=black or black british, 5=chinese, O=other)

Table 5.4: Sampling Grid Y5+Y6 Firs Hill n=5

Health Level	M1	M2	M3	M4	M5	MO	F1	F2	F3	F4	F5	FO
1									1			
2			1				1		1			
3			1									
4												
5												

(M=male, F=female, 1=white, 2=mixed/dual heritage, 3=asian or asian british, 4=black or black british, 5=chinese, O=other)

Table 5.5: Sampling Grid Y3+Y4 Hunters Bar n=24

Health Level	M1	M2	M3	M4	M5	MO	F1	F2	F3	F4	F5	FO
1	3	1	1				3		1			
2	2	1				2	1	2	1			
3	2						2					

4									1			
5							1					

(M=male, F=female, 1=white, 2=mixed/dual heritage, 3=asian or asian british, 4=black or black british, 5=chinese, O=other)

Table 5.6: Sampling Grid Y5+Y6 Hunters Bar n=28

Health Level	M1	M2	M3	M4	M5	MO	F1	F2	F3	F4	F5	FO
1	3						1	3		1		2
2	3		1	1			3	2				
3	2	1	1				2	1				
4	1											
5												

(M=male, F=female, 1=white, 2=mixed/dual heritage, 3=asian or asian british, 4=black or black british, 5=chinese, O=other)

Table 5.7 below shows a summary of the characteristics of the two age groups. A good balance of gender and ethnicity was achieved, however there were few children with fair or poor health due to a lack of children with these levels of health in the sample. χ^2 tests or Fischer's exact tests were undertaken (as appropriate given the number in each category) to compare the groups and found no difference between gender, level of health and ethnicity between the groups (at $p<0.05$).

Table 5.7: Characteristics of the Samples

Characteristic	n (7-9 years)	n (9-11 years)
Total sample	41	33
Male	19	15
Female	22	18
Y3	22	-
Y4	19	-
Y5	-	18
Y6	-	15
Excellent health	13	11
Very good health	14	13
Good health	10	8

Fair health	2	1
Poor health	2	0
White	8	16
Mixed/dual heritage	2	7
Asian or Asian British	5	6
Black or Black British	1	2
Chinese	0	0

Children concentrated well during the interviews, the warm up exercise worked very well and children seemed to feel relaxed and at ease with the interviewer. The consent process was very helpful for explaining to children the purpose of the research and what was involved and some children raised questions prompted by this. Some children commented that they had enjoyed the interview. No children asked to stop the interview and no bad feedback or experiences were reported. Interviews varied in length from 4 to 26 minutes.

Saturation was reached in both age groups. Despite increasing the number of interviews and specifically trying to target children with poorer health, no new themes or issues emerged. The total number of interviews was more than needed, but this has increased the confidence that saturation was reached. Saturation was reached at around about 30 interviews in each group.

A range of health problems arose in the interviews, covering both chronic conditions such as asthma, epilepsy, allergies and eczema, and acute conditions such as flu, infections and headaches. Some were minor and some more serious, for example some cases of eczema required hospital treatment. Some conditions also arose that involved hospitalisation, including pneumonia and muscle growth problems. Many conditions required treatment via the GP, such as hay fever, ear infections and chicken pox and some required hospital visits, such as losing feeling in the legs, severe asthma and broken bones. A list of health problems by age group is given in Appendix 5.G and 5.H.

The frameworks for both age groups are given in Appendixes 5.I and 5.J. The analysis generated ten dimensions for each age group which are broadly similar, with the exception of the final dimension. They are listed in Table 5.8 below. The results reported here are from both age groups.

Table 5.8: Dimensions of HRQoL

Dimension	7-9 Years	9-11 Years
1	Worried Scared	Worried
2	Sad Upset	Sad Upset Unhappy Miserable
3	Annoyed Frustrated	Annoyed Frustrated Angry
4	Hurt Pain	Hurt Pain
5	School work	Learning
6	Daily Routine	Daily Routine
7	Tired Weak Drowsy	Tired Weak Energy Weary
8	Joining in activities that want to	Joining in activities that want to
9	Sleep	Sleep
10		Embarrassed
11	Jealous	

<u>1. Worrying/Scared</u>

This dimension stemmed from feelings of worry about their health or illness and what was going to happen to them, both in the short and long term, for example worrying about whether their health would get worse and worrying that they would always have the health problem. There was also worry because of the physical symptoms, for example when breathing was restricted due to an asthma attack or bleeding from eczema. Some children worried about what was going to happen to them, for example if they had an allergy, they were worried about what would happen if they had a reaction. The younger children tended to talk more about being scared or nervous, sometimes from people looking at them because of their health and they were also scared because of the physical effects of their health. Older children just talked about it in terms of being worried. For example:

Child: Sometimes it's hard to breathe, when you're breathing up or down sometimes a little ... (unclear)... sometimes when I get tired it's hard.

Interviewer: It's hard to breathe, how does that make you feel?

Child: A bit worried.

Interviewer: A bit worried, why do you get worried?

Child: Cos sometimes like err you don't know what to do.
(B127, male, 11, very good health)

Child: Because sometimes I worry what's going to happen to me.
(B12, male, 8, very good health)

Child: I felt like really, really worried, worry me, like, um it felt really scratchy and itchy just really worried like it's never gonna stop and it's never gonna go away.
(B33, female, 8, fair health)

Interviewer: How does it make you feel when you have eczema?

Child: A bit worried because like I get worried in case, in case, because once like I used to scratch all here and I used to start bleeding, last

night I got really worried in case if all my, if all my body gets, if all my body bleeds, I just get really worried like that.
(B33, female, 8, fair health)

2. Feeling sad or unhappy

There were several reasons for feeling sad, miserable, upset or unhappy. Some children said it was because they felt unwell and experiencing the physical symptoms made them feel sad. Some children felt sad when they couldn't do things they would normally do, like going to school, seeing their friends or doing activities they normally did. Some children felt sad when they were teased because of their health. All these reasons occurred in both age groups and the language was the same apart from older children using the term miserable in addition to sad, unhappy and upset.

Child: I just felt unhappy because I'm missing things that I normally do on the weekend.
(A22, female, 10, excellent health)

Child: Yeh it stopped me skipping cos I can't skip with one arm.

Interviewer No, how did that make you feel?

Child: Upset because I like skipping.
(A8, female, 8, good health)

Interviewer: And can you tell me about that, what's it feel like, when you get the chest pains.

Child: It just gets all squeezed up and I just keep on rubbing it and it just feels, it just feels hurt and I don't want it to happen but it just does happen.

Interviewer: And how does it make you feel?

Child: Unhappy, actually.

Interviewer: Unhappy, why does it make you feel unhappy?

Child: Because it hurts and I don't want it to hurt actually.
(A13, female, 9, good health)

3. Feeling angry, annoyed or frustrated

This dimension arose from a variety of reasons for feeling angry, annoyed or frustrated. In some cases, physical symptoms caused children to feel annoyed or frustrated. Some children also got annoyed because their health affected their everyday activities or things that they did, including sleeping, eating and being able to concentrate. Some younger children also mentioned being annoyed because of having to apply medication and also having to go to hospital for tests.

Interviewer: You couldn't breathe. And what does that feel like when you can't breathe?

Child: I get really frustrated.

Interviewer: Why is it frustrating?

Child: Because you wanna breathe and I can't breathe.

(A17, male, 9, very good health)

Child: I had to get into the bath and put a glove on and then I had to put a plastic bag on and then you've got soap and I it's just really annoying because I couldn't really do very much and, then I couldn't play football and then like I couldn't do very much stuff basically and I didn't like it.
(B97, female, 10, excellent health)

Child: It was really annoying I couldn't go to sleep at night, cos I was coughing.
(B3, male, 8, good health)

4. Hurting/pain

Many different health problems led to pain of different degrees. Sometimes pain came about through itchy skin, for example, when talking about spots on their skin. In other cases pain came about through physical symptoms like a sore throat, an asthma attack, tummy ache, headaches or coughing. Other children talked about the pain of treatments they have to have in hospital. Others talked of pain when they undertook specific activities, such as pain when they walked on their sore feet. The same reasons and issues came up in both age groups and both age groups used the same terminology, describing it as hurt or pain.

Child: Umm it's like whenever you swallow there's like it's horrible it's like a dry but sharp pain
(B106, female, 11, excellent health)

Child: It's like, it was like I couldn't breathe properly and when I tried to breathe in it really hurt my chest.

(B110, female, 11, good health)

5. Learning/schoolwork

There were many issues to do with school and they all led to the same outcome, in that problems around concentration, being absent from school because of health and not being able to manage work, led to schoolwork and learning being affected. Some children said that their physical symptoms meant they were not able to concentrate, which meant they could not learn. This included pain or itching or being tired. Other children said that when they were off school because of their health, it affected their learning as they were missing their lessons, or in some cases it was because they were unwell at school and had to miss lessons. Some children described how problems with vision, hearing and speech all led to them having difficulty in lessons, so it affected their schoolwork, for example not being able to hear the teacher or see the board properly. The same issues arose for both age groups, however the younger children talked about their work or school work, often naming specific parts of it such as writing or drawing, whereas the older children tended to extend this by saying it affected their learning in general.

Child: I'd just sit down and be coughing a lot and disturb the class, I wouldn't be learning because I'd be concentrating on my cough instead of learning.
(A18, male, 10, good health)

Child: cos I can't concentrate on my work its like, it all just goes all over blurry so I close my eyes and then I look back and it goes ok but then like it hurts my eyes.
(B97, female, 10, excellent health)

Child: It affects me kinda like speaking because all your mouth dries up, dries up and you can't really open your mouth.

Interviewer: And what does that affect when you can't speak, how does that affect you?

Child: I can't answer any answers at school or anything.
(B7, female, 8, excellent health)

Interviewer: How does it affect your work?

Child: Cos I can't see and I don't know what to do.
(A9, male, 8, very good health)

6. Daily routine (eating, bathing, dressing, getting ready, moving around)

There were many issues around children's daily routine, including being able to have a bath and wash themselves, being able to get dressed and get ready for school. Some children found it took longer to get ready for school as they had to take their medication. Some children had problems undertaking other daily tasks including picking things up, and getting out of bed. In addition, issues arose around not being able to eat, for example if children had a sore throat or stomach ache, then they would not be able to eat what they normally would. Some children did not eat so they wouldn't be sick and some children lost their appetite when their health was not good.

Some children had problems moving around, for example going up and down stairs or getting around places. The issues were the same in both age groups. Neither age group talked about a daily routine, but instead talked about the individual tasks that they would normally have to do as part of their everyday life. The younger children sometimes broke it down into smaller tasks, for example specifically talking about difficulties with cleaning their teeth, rather than a more general getting ready in the morning.

Interviewer: How did it affect you when you had your leg like that in the things that you do everyday?

Child: Well it's hard to get up the stairs, that was the main thing so I had to crawl up the stairs but...
(B122, male, 10, good health)

Interviewer: And how does it affect you at home?

Child: Umm it takes a bit longer to get ready for school.

Interviewer: Why's that?

Child: Erm, because I've got lots of things to do to get ready I don't have to just brush my teeth and then that's it I have to my (unclear) brush my teeth and then I have to do this special thing with my teeth.
(B70, male, 10, very good health)

Child: Getting dressed was quite hard.

Interviewer Was it?

Child: Yeh, cos I had to stretch my arm a bit to get something on, like a jumper, I had to stretch my arm a bit and then it hurted.
(A14, male, 9, good health)

Child: well I couldn't walk, I don't know why but my legs I couldn't move it so my mum got me up and made me stay in bed all for two weeks without moving.
(B23, male, 8, very good health)

7. Feeling weak/weary/tired

Issues arose around feeling weak and not having any energy to do things, Some children didn't want to do anything because they had no energy, or because of how they felt when they had symptoms, for example, a headache. Some children felt tired and for different reasons, sometimes it was because their health problem made them feel like going to sleep, sometimes it was because their sleep had been affected. Some children had to rest more or sleep more. Older children talked more about having enough energy to do things and feeling weary. Younger children sometimes described feeling drowsy.

Child: I felt really weak so I couldn't really do a lot of activities that took lots of energy cos I didn't really have a lot of energy at that time.
(B110, female, 11, good health)

Child: Well it was like a fluey sort of cold or something, I was really tired and I just fell asleep on the couch. I just sort of felt ill and just stayed on the couch for a week.
(B120, male, 11, excellent health)

Child: Err, well I usually have to have like a rest on the sofa cos it feels really bad and tiring.
(B22, female, 8, good health)

8. Able to join in activities that they want to (e.g. playing out with friends, sports)

There were a lot of issues that arose around being able to join in activities that children wanted to do. These activities included playing with or being with their friends, playing out, going on trips, or joining in activities or sports that they wanted to. In some cases children could still play out, but their health made it difficult. Some children had to miss out on sports because they had hurt themselves or they were in pain. In some cases children could not do physical activity as it set off their illness. In some cases they could still join in but found it harder. Some children said they could not play much because they were in hospital. Other children described how they missed their friends when they were poorly. In some cases this was because they were off school or in hospital. The issues were the same for both age groups.

Child: It affects me because at home I can't do nothing cos I sometimes I like playing with my brother's and sisters on the road or outside and I have to stay inside and do nothing.
(A18, male, 10, good health)

Child: Yeah I wasn't able to join in with things that I normally do like tennis and football and stuff like that.
(B114, male, 11, very good health)

Child: I didn't get to meet with some of my best friends and play games with them.
(B42, male, 9, very good health)

9. Sleep

Sleep emerged as an issue among children for different reasons. Some children found it difficult to get to sleep, some could not sleep because they were worried, others because of symptoms, such as coughing or being sick. Some children experienced broken sleep, in some cases this would really affect their sleep as they could not get back to sleep and in other cases not so much. Some children really struggled with sleep when all they wanted to do was go to sleep and other children hated going to bed because they knew they would not be able to sleep and would have to lie there alone. This had consequences for the next day, when children felt tired and found it difficult to get up for school and concentrate. Some children could not sleep at all and some children woke early. These issues arose in both age groups and were described in very similar ways.

Child: Oh that was horrible, it was absolutely horrible, you couldn't go to sleep at all because when, when...cos if you opened your mouth and tried breathing out of your mouth you swallowed because my mouth would be all dry in the middle of the night and I'd have to swallow and it wasn't very nice
(B106, female, 11, excellent health)

Child: Yeah I couldn't sleep cos it really hurt my throat when I slept, so I couldn't sleep at all.
(B110, female, 11, good health)

Child: In the night I wake up because I'm scratching it.

Interviewer: And how does that feel?

Child: Horrible because I can't get back to sleep.
(B28, male, 9, excellent health)

Child: when um I kept being sick in the night and then, um, I didn't get much sleep then cos I just had to kept waking up and stuff yeh.

(A15, female, 9, good health)

10. Embarrassed

This dimension arose in the older children. Some children said they were teased because of their appearance, in some cases because of their figure. In some cases, children were embarrassed about their appearance and took action to cover it up. Children did not like crying or being sick at school because they felt embarrassed. Children also talked about how their appearance bothered them because of their health, for example because of having to wear glasses, or when they had visible health symptoms (for example they were embarrassed when other people saw their eczema) or had to take medication.

Child: but because of my figure, but, because of my figure, because my bones are quite heavy so, I'm quite big so sometimes I get a bit teased about that
(A20,female, 11, very good health)

Child: cos everyone sees you in the class and you feel a bit embarrassed.
(B108, female, 10, very good health)

Child: sometimes I forget to cream my legs and I get a bit embarrassed in P.E.
(B82, female, 10, good health)

11. Feeling jealous

This arose in the younger children and was similar in some ways to the embarrassed dimension that arose in the older group, however it was not as strong. Some of the issues children raised were hinting at being embarrassed.

Children felt jealous of others as they did not want to have their health problem and wanted to be like others who did not have any problems. They also felt jealous when they could not join in activities that others could because of their health.

Child: because like I think other peoples legs and arms are not like that I wish mine weren't like that and things like that.

Interviewer: How does it make you feel?

Child: Umm a bit jealous because like, people, um other people don't have it on their body, I wish I was the same as them because I'm like the only person who has eczema on my body in my family and be like thinking like well how come she's got it.
(B33, female, 8, fair health)

Child: sometimes I feel I wish I didn't wear them, sometimes yeah
(A15, female, 9, good health)

One overall theme that arose was the idea of being normal, or children thinking of what a usual role would be for a child of their age and then thinking how they differed from this, for example, saying they could not do what they would normally do when they were well, or could not do what other children their age do. They also sometimes expressed a desire to be like other children who did not have the health problems they had or to be what they described as normal, meaning free of the health problem.

5.5 Discussion

The interviews worked best in a one to one setting as sometimes children would be discussing sensitive information and indeed 73 out of 75 children chose this format. The format of the interview worked well as children first became comfortable with providing descriptive information about their health and then went on to talk about how it affected them in their lives. This research has demonstrated that it is feasible to interview children about their health and that they are able to understand and describe how health affects their lives.

This qualitative study has generated a wide range of dimensions of HRQoL. There were many reasons why the dimensions emerged, as different health problems could result in the same impact on HRQoL. For example, if a child's hearing was affected, it sometimes meant they found lessons difficult as they could not hear the teacher and it also meant they had some pain in their ear. Other children had an allergy which affected their lessons because their concentration was affected and they were scratching which also gave rise to pain. Also, breathlessness led to children feeling worried and frustrated.

Most dimensions arose from children across the range of levels of health and through both acute and chronic health problems. However, feeling scared or worried seemed to arise mainly in children who had chronic health problems such as asthma or eczema who worried about when their symptoms came back and what would happen to them. Embarrassment arose mainly through health problems which gave rise to visible symptoms such as rashes, or for children who felt they were overweight. Feeling sad or worried did not arise as health problems in their own right, rather they arose as a consequence of other health problems.

Quite often, the way in which children thought about these dimensions was in terms of the overall theme of a departure from normal, meaning children were thinking of what a usual role would be for a child of their age and then thinking how they differed from this.

This work has some evidence that there is a similar HRQoL framework for 7-9 and 9-11 year old children as the same dimensions (except one) arose from each group. Even though the language and terminology used by the age groups sometimes differed, they were describing the same concepts. This gives confidence that one measure can be developed for both groups, however more work is required on the appropriate wording to use, to make sure it is suitable for all ages.

The dimensions identified include many of the areas of HRQoL covered in widely used adult PBMs, such as the SF-6D (Brazier 2002) and the EQ-5D (Dolan 1997), such as physical, emotional and social aspects. For example, the EQ-5D contains a pain dimension and an anxiety/depression dimension. Whilst not described in the same terms, the worrying and sad/unhappy dimensions in this work are similar in concept to the latter. Similarly the SF-6D contains dimensions on pain, mental health, vitality, social functioning, physical functioning and role limitations. Some of the dimensions developed in this work, such as feeling tired or weak and joining in activities, seem similar in concept to vitality and social and physical functioning. The dimensions also broadly conform to the World Health Organization (WHO) classification of health (World Health Organisation); that health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

Although the dimensions are intended to be mutually exclusive, there are relationships between them, for example, not being able to join in activities may make a child feel sad, or being in pain may mean that a child may not be able to sleep. This is perfectly acceptable for a HRQoL measure as it is often the case that people have problems on several dimensions at once, however this may have implications for the design of the valuation survey. Although the dimensions may be related through these problems, they are still different dimensions of HRQoL and reflect the differing impact of problems on children's lives.

From the literature review in this thesis, it was found that the PEDSQOL(Varni 1999), CHQ (CHQ)(Eiser 2001) and the HUI2(Torrance 1996) were the most frequently used generic paediatric instruments. Table 5.9 below lists the dimensions of these, together with those of the new measure.

Table 5.9 Comparison of dimensions with three existing paediatric generic instruments

New measure	HUI2	PEDSQoL	CHQ (CF87)
Worried/Scared	Emotion	Worry**	Mental health
Sad/Upset/Unhappy/Miserable		Psychological* functioning	Role/social functioning (emotional)
Annoyed/Frustrated/Angry			Self-esteem
Jealous			
Embarrassed			
Tired/Weak/Drowsy/Energy/ Weary			
Hurt/Pain	Pain	Pain**	Bodily pain
			Discomfort
Sleep			
Daily Routine	Self Care		
School work/Learning	Cognition	Cognitive problems**	
Joining in activities that want to	Mobility	Social functioning*	Physical functioning
		Physical functioning*	Role/social functioning (physical)
	Sensation		
		Physician/nurse communication**	
		Perceived physical appearance**	
		Nausea**	
		Treatment anxiety**	
		Procedural anxiety**	
			Parental impact (emotional)
			Parental impact (time)
			Family activities
			Family cohesion
			Global item: change in health
			General health perceptions
			General behaviour

*Core measure
**Symptom/treatment related module

It shows some similarity in coverage but also some important differences. The HUI 2 (Torrance 1996) is the only existing paediatric generic PBM. It measures impairments and takes a “within skin” approach rather than assessing the impact on function. For example, sensation is derived from questions asking about vision, hearing and speech. The dimensions developed in this study go beyond this, in that the consequences of hearing and vision problems were investigated, for example in joining in activities or being able to manage their school work. Whilst there are some similarities to the dimensions developed here, for example *pain* and *daily routine*, there are many differences, for example the HUI2 does not include *sleep, jealous, embarrassed* and *tired/weak*.

The PEDSQoL (Varni 1999) goes beyond how health affects children in that it includes areas such as *physician/nurse communication*. The CHQ (Eiser 2001) includes more than HRQoL with dimensions related to the family, for example *family activities* and *family cohesion* and also includes a *behaviour* dimension. There are also some similarities with the dimensions developed here, for example *pain* and *emotions*. *Sleep* and *feeling tired/weak/having no energy* is a gap across all the instruments and is a key difference in the measure developed here.

The qualitative approach taken here of directly looking for dimensions and the explanations behind them is similar to that taken by Grewal (2006) in their development of a generic QoL measure for older people. In this work, they conceptually grouped data into mutually exclusive attributes of QoL. It is in contrast to the more common approach taken in the instrument development literature, where large lists of items are generated by the developer from the literature and opinion and then a technique such as factor analysis is used to develop dimensions (Drotar 2006) (Riley 2004). A qualitative approach can also be used to generate this long list of items and then factor analysis used to develop dimensions, however the key difference in the qualitative approach adopted here, generating dimensions directly, is that it is more useful in developing a PBM since it generates one item per dimension and levels within these in order to construct a health state classification. There is also a stronger use and reliance on psychometric techniques in the item generation approach.

Involving children in developing the content of the descriptive system helps ensure content and face validity as it should be comprehensive in its coverage and be appropriate for the population. This work is also in line with the Food and Drug Administration guidelines on patient-reported outcome measures, which recommend that development of paediatric instruments should consider age-related vocabulary and language comprehension, and fairly narrow age groupings should be used to account for developmental differences (Food and Drug Administration).

There was breadth of coverage in terms of acute and chronic health problems in the sample, although it is acknowledged that there are gaps at the lower end, in that there were comparatively fewer interviews carried out on children whose health was rated as fair or poor. This is not surprising given that the population was recruited through schools and so tended to exclude those children who are in hospital or not in mainstream education. Sampling in the community via schools meant that the majority of child health problems were covered and only the very sick (e.g. terminally ill children or those in hospital permanently) were not included. It would be very difficult ethically to interview children such as these about how their health affects their lives at such a young age (7-11 years) when they are so ill. A wide range of acute and chronic health problems were included however and so it was felt the sampling frame was sufficient. For example, there were children who had been in hospital many times for many kinds of conditions, including pneumonia, having no feeling in their legs, muscles not growing properly in the stomach and epilepsy/fits.

The use of the rating of health (Excellent, very good, good, fair, poor) by the parent was only used as a proxy for sampling children with a range of health problems and there were actually many children who had quite bad health problems whose parents had rated their health as excellent.

There was a notable difference in response rates between the schools. There could be many reasons for this, including the fact that Firs Hill school has 67% of pupils with a first language other than English and in many cases, the parents do not speak English, and therefore may not have read/understood the letter and information leaflet asking for consent to approach their child. Consent or response rates vary widely in health research (both for adults and children) and often depend on the nature of the research being carried out, for example what it would involve for the child and how long it would take. However, school based research usually yields a higher consent rate than research conducted in hospitals or a home setting.

It is acknowledged that the population interviewed in this research all come from Sheffield and it is not certain that the results of this will apply nationally. Compared to National data, the percentage eligible for free school meals and the total percentage of children with special educational needs is higher for Firs Hill School and lower for Hunters Bar, however there are more children from ethnic minorities and with a first language other than English in both schools than nationally. The breadth of socio economic diversity reached in the sampling should ensure that the measure is likely to be applicable more widely than Sheffield and this is something that can be tested in future research.

These dimensions are intended to form a generic measure and need to be suitable for a wide range of conditions and health problems found in the general paediatric population. It may be that the measure does not perform so well in more serious paediatric health conditions, as the measure has been developed on a relatively healthy sample. Testing the measure on clinical populations will help to determine how well the measure performs in different patient groups.

5.6 Conclusion

This research has identified the ways in which children age 7-11 years say health affects their lives. A range of dimensions of HRQoL emerged, covering physical, social and emotional aspects. In contrast to existing paediatric measures, the dimensions have been developed directly from the relevant population, increasing the content and face validity. There was also no influence from any other source, such as parents, teachers, medical professionals or the literature as bottom up methods were used. This research has also demonstrated that children of this age are able to provide information on HRQoL and gives some evidence of a common framework across age. The next stage in the research was to develop a descriptive system that is amenable to valuation and to test it in different paediatric patient populations.

Appendix 5.A: Topic guide used in the piloting work

Topic Guide for Interviews with Children

Objective
To identify the dimensions of quality of life that matter to children that are related to their health

1. Introduction

- Introduce myself
- What we will be doing – health and how it affects our lives
- Opinions – no right or wrong answers
- Free to stop at any time (agree signal)
- Names - Badges/Stickers
- Explain about the microphone/confidentiality

2. What you think

Health

- Tell me about your health

3. How does your health affect your life?

- Think about the things that you can or can't do because of your health
    - Probe: home, school, other places (community)
    - Probe: family, friends, weekends, spare time

Appendix 5.B: Letter to Parents

Hunter's Bar
JUNIOR SCHOOL

09 December 2008

Dear Parent/Guardian,

Research to develop a new health questionnaire for children

I am a researcher at The University of Sheffield and I am undertaking research over the next 18 months to find out what children think about how health impacts on daily life. This important and new research will give us the chance to find out what children think.

I will be interviewing about 50 children and it is important to include a mixture of different race, gender and experiences of health.

I am writing to ask for your consent for your child to be asked to take part. If you are happy for them to be asked, please complete and sign the enclosed consent form and return it to the class teacher. This information is strictly confidential and will be used only by me for sampling purposes.

I have enclosed an information sheet which describes the research and should help to answer many of the questions you may have.

If you have any further questions, please do not hesitate to contact me on the details below.
Many thanks for your cooperation.

Yours sincerely

KS Stevens

Katherine Stevens
The University of Sheffield

Direct telephone: 0114 2220841 (please leave a message if I am not there)
E mail: K.Stevens@Sheffield.ac.uk

Appendix 5.C: Information Leaflet for Parents

The
University
Of
Sheffield.

INFORMATION SHEET FOR PARENTS/GUARDIANS

Research to develop a new health questionnaire for children

What is the purpose of the research?
Society has to make choices and decisions in the National Health Service (NHS) and it is important we do this as well as we can. The aim of this study is to develop a health questionnaire for children, to inform decisions in the NHS. I will be working with a range of children in order to develop this questionnaire. I want to find out what children think is important in health.

Why is this research necessary?
At the moment, there are very few questionnaires suitable for children. Research in the past has used adult measures on children, but they are not really suitable. This research aims to create a new questionnaire for children. This will lead to making better decisions in the NHS. I feel it is important that children's views are taken into account.

Why has my child been chosen?
Your child's school has agreed to take part in the research. I am writing to the parents of all children in the school to ask for consent to ask their child to take part. I am also asking for information about your child's level of health so that I ensure I ask a range of children with different experiences of health. I will choose children from a range of backgrounds to make sure the research is representative. Your child may or may not get asked to participate.

Do they have to take part?
They do not have to take part. If you don't want them to, you need not do anything, I will not ask your child without your written consent. Children may also choose not to participate at any stage of the research.

What does it involve?
There are three parts. Your child may be asked to take part in one or more of them. The first is a small group interview of approximately 20-30 minutes. Children will be asked to think about the things that are important to them in their everyday life. The interviews will be carried out in small groups so that children feel comfortable. It will be made clear that it is not a test and there are no right or wrong answers, I just want to know what they think. The interviews will be recorded but all children will remain anonymous. Children can choose to do the interviews on their own if they prefer.

The second part will take all the ideas and opinions that children come up with in the interviews and ask them to choose which are the most important. This second stage will be carried out once all the initial interviews are finished. This second part should not take longer than 15 minutes. The children will do this task by themselves. Again, there are no right or wrong answers, I just want to know what they think. What the children select as the most important will be used to make the questionnaire.

The third part will ask children to try out the questionnaire and see what they think about it. It should not take very long and they will complete it by themselves. Their answers will be confidential.

Is the research confidential?
All the research is strictly confidential and your child's identity will not be revealed to anyone. They will remain anonymous in all analysis and reporting of the research.
The only exception to this is if they disclose any information which raises child protection concerns, in which case the information will be passed on to the school's protection officer in line with the school's policy and the National Children's Bureau Guidelines for Research.

Who will have access to the data and where will it be held?
All data will be held in confidence at The University of Sheffield under the control of me (Katherine Stevens). It will be used only for the purposes of this research and not passed on to anyone else. The two supervisors of the research (Professor Chris McCabe and Professor John Brazier) and any peer reviewers will have access to the anonymised data, under my control. This is because they may want to check the analysis.

When and where will the research take place?
The interviews will take place at your child's school in a quiet space but where other adults can observe. I will conduct the interviews myself. Your child will not be alone with the researcher and in accordance with the school's policy I have enhanced CRB checks. Every care will be taken not to disrupt normal lessons.

What if I change my mind?
You may change your mind at any point, the children will not be affected in any way and they will no longer take part in the research. If you do change your mind, please contact me on the details below.

What if my child changes their mind?
Children are free to change their mind at any point. This is up to them and they will be returned to their class if they do, and carry on as normal.

What will happen to the results of this research?
The results of this research will be used to develop a health questionnaire for children. It is hoped that the questionnaire will be useful for health researchers in the future in making decisions about the healthcare of children. The results will be presented both nationally and internationally and I will keep the school informed.

Who is funding the research?
This research is part of a Special Training Fellowship in Health Services and Health of the Public Research, funded by the UK Medical Research Council. (www.mrc.ac.uk)

Has the research been approved by an ethics committee?
The research has been approved by the University of Sheffield Ethics Committee. It also has the support of the Children and Young People's Directorate at Sheffield City Council.

Who should I contact for further information?
If you have any questions about the study or require any further information, please do not hesitate to contact me on the details below:

☎Direct telephone:	0114 222 0841 (please leave a message if I am not there)
🖳E mail:	K.Stevens@Sheffield.ac.uk
🖃Post:	Katherine Stevens Health Economics and Decision Science, ScHARR The University of Sheffield Regent Court, 30 Regent Street Sheffield, S1 4DA

THANK YOU FOR TAKING THE TIME TO CONSIDER THIS RESEARCH

Appendix 5.D: Consent Form

The
University
Of
Sheffield.

PARENT/GUARDIAN CONSENT FORM

Title of Project: Research to develop a new health questionnaire for children

Name of Researcher: Katherine Stevens, The University of Sheffield, UK

Please tick (☑)

1. I confirm that I have read and understood the information sheet for parents and have had an opportunity to ask questions. ☐

2. I understand that my child's participation is voluntary and that they are free to withdraw at anytime at my or their request, without giving a reason. ☐

3. I agree that my child can be asked if they would like to take part in the above study ☐

Name of child:________________________________

Age of child:________________________________

Name of parent/guardian:________________________

Signature of parent/guardian:______________________

Today's date:________________________________

Please answer the following questions by ticking ONE box only, like this ☑

Overall, how would you rate your child's health during the past week?

☐ Excellent
☐ Very good
☐ Good
☐ Fair
☐ Poor

Gender of your child:

☐ Male
☐ Female

Ethnic Origin of your child:

☐ White
☐ Mixed/Dual heritage
☐ Asian or Asian British
☐ Black or Black British
☐ Chinese
☐ Other ethnic group (please specify)________________

ALL ANSWERS WILL REMAIN STRICTLY CONFIDENTIAL

Thank you. When completed, please return this form to your child's class teacher.

Any questions?

Please feel free to contact me by:

☎ Telephone: 0114 222 0841 (please leave a message if I am not there)

💻 e mail: K.Stevens@Sheffield.ac.uk

🖃 Post: Katherine Stevens
Health Economics and Decision Science
ScHARR
The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA

Appendix 5.E: Topic guide used in the main interview work

Topic Guide for Interviews with Children

Objective
To identify the dimensions of quality of life that matter to children that are related to their health

1. Introduction

- Introduce myself
- What we will be doing – health and how it affects our lives
- Opinions – no right or wrong answers
- Free to stop at any time (agree signal)
- Names - Badges/Stickers
- Explain about the microphone/confidentiality

2. What you think

Health

- Tell me about your health (make sure acute and chronic both recognised as valid – probe to check)

3. How does health affect your life?

- Think about the things that you can or can't do because of your health
    - o Probe: home, school, other places (community)
    - o Probe: family, friends, weekends, spare time

Appendix 5.F: The process of using framework.

The following steps are the process of Framework (Ritchie 2005) (numbered) and under each in italics is the process specific to this research.

1. Review data/familiarise

   Reread and relistened to all interviews
2. Identify recurring themes or ideas

   Noted down as they occurred
3. Devise conceptual framework/index using these themes and topic guide

   Grouped themes together
4. Themes lead to a smaller number of main themes so you have key themes and sub topics.

   Developed main themes and sub themes (some themes collapsed into one, as the same/very similar) to give thematic framework.
5. Indexing/coding of data according to these. (adapt framework accordingly if necessary)

   All data coded in NVIVO. A tree node was a theme and a free node was a sub theme.
6. Charting (summarizing/synthesizing the data within a matrix)

   Each sub theme was recalled in NVIVO and each reference related to the sub theme was charted. Terminology was kept the same. Italics represented verbatim quotes, standard text represented a summary. Data were synthesized and summarized.
7. Each theme has a chart with the sub themes within it.

   All data was charted in this way. 7 themes plus sub themes within each were developed.
8. Abstraction

   Each sub theme was reviewed and explanations behind the affected areas of HRQoL were looked for. Relationships between themes and subthemes were mapped. Generated a higher order of dimensions. Generally a health problem gave rise to several symptoms/physical *characteristics, e.g. my throat was sore. What we are interested in is how this affects them. E.g. I find it difficult to eat anything and my throat hurts. E.g. 2 hearing was affected – this meant that they found school difficult as they couldn't hear the teacher in the lessons.*

Appendix 5.G: Y3 and Y4 List of Health Problems Covered in the Interviews (n in brackets)

Headache (9)	Allergies – various (3)
Feeling sick, being sick (11)	Nose bleeds (2)
Hearing, glue ear – grommets (2)	Ear ache (3)
Poor vision (1)	Muscle not growing properly in stomach (1)
Tummy ache (10)	Fever (1)
Eye infection (1)	No feeling in legs (1)
Dyslexia (1)	Badly cut nose (1)
Asthma (9)	Nausea (1)
Broken arm (2)	Verruca (1)
Tooth decay (1)	Chicken pox (1)
Leaky ear (1)	Heat rashes (1)
Flu (2)	Sensitive to food colouring/hyperactivity (1)
Pneumonia (1)	Tonsillitis (3)
Hay fever (6)	Sticky/lumpy eyes (1)
Cough (3)	Twisted ankle (4)
Spots/rash (2)	Fits (1)
Sore throat (4)	Eczema (10)
Broken toe (1)	Itchy eyes (1)

Appendix 5.H: Y5 and Y6 List of Health Problems Covered in Interviews (n in brackets)

Eczema (various degrees of severity) (6)	Blocked nose (3)
Asthma (various degrees of severity) (10)	Rash (1)
Flu (2)	Blocked up ears (1)
Hay Fever (various degrees of severity) (6)	Tiredness (1)
Headaches (various degrees of severity) (8)	Heat rashes (1)
Sick (feeling and being) (9)	Tonsillitis (5)
Nose bleeds (1)	Stress (1)
Overweight (2)	Throat infection (1)
Cough (chesty, tickly) (3)	Weak wrists and ankles (1)
Dental health (3)	Broken finger (1)
Sore throat (3)	Cramp (1)
Cold (various degrees of severity) (9)	Not able to sleep (various reasons) (2)
Chicken Pox (3)	Broken wrist (2)
Molluscus/Spots (3)	Grommets in ears/hearing (2)
Allergies – various, including animals, washing up liquid, chlorine (3)	Knee ligament damage (1)
Stomach ache (4)	Twisted and bruised ankle (1)
Ear ache (2)	Stomach bug (1)
Sprained ankle (1)	

Appendix 5.I: Y3/Y4 Framework

School/Education

- Concentration
- Off school
- Managing work

Emotions

- Worried/nervous
- Bored
- Unhappy/sad
- Annoyed/irritated
- Angry/jealous

Physical Activities

- Sports/games
- Going places and doing things
- Playing outside

Social

- Appearance
- Friends
- Joining in

Senses

- Hearing
- Talking
- Seeing

Sleep

- Sleep

Physical Feelings

- Pain/hurting
- Tired/weak
- Feeling unwell

Daily activities/routine

- Moving around
- Everyday tasks

Appendix 5.J: Y5/Y6 Framework

School/Education

- Concentration
- Off school
- Managing work
- Learning/education

Physical

- Pain
- Sick
- Tired
- Energy/feeling weak
- Feeling unwell/poorly
- Not wanting to do anything
- Sleep

Emotions

- Anxious/worried
- Isolated/bored
- Unhappy
- Annoyed

Physical Activities

- Sports/games
- Bathing/self care
- Playing out

Social

- Appearance
- Teased
- Friends

- Joining in

Senses

- Hearing
- Talking
- Taste
- Smell

Food

- Eating

References

Brazier, J.E., Roberts, J. & Deverill, M. 2002. The estimation of a preference based measure of health from the SF-36. Journal Health Economics. 21(2), 271-292.

Dolan, P. 1997. Modelling valuations for EuroQol Health States. Medical Care. 35, 1095-1108.

Drotar, D., Schwartz, L., Palermo, T. M., & Burant, C. 2006. Factor Structure of the Child Health Questionnaire-Parent Form in Pediatric Populations. Journal of Pediatric Psychology. 31(2):127-138.

Eiser, C. & Morse, R. 2001. Quality-of-life measures in chronic diseases of childhood. Health Technology Assessment. 5(4).

Food and Drug Administration, U.S. Department of Health and Human Services. February 2006. Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Draft Guidance.

Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J. & Coast, J. 2006. Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? Social Science and Medicine. 62: 1891-1901.

NVIVO7 QSR International

Riley, A.W. 2004. Evidence that school-age children can self-report on their health. Ambulatory Pediatrics. 4(4):371-376.

Ritchie, L., Lewis, J. (Eds.) 2005. Qualitative Research Practice; A Guide for Social Science Students and Researchers. SAGE publications, London.

Torrance, G.W., Feeny, D.H., Furlong, W.J., Barr, R.D., Zhang, Y. & Wang, Q. 1996. A multi-attribute utility function for a comprehensive health status classification system: Health Utilities Mark 2. Medical Care. 34(7):702-722.

Varni, J.W., Seid, M. & Rode, C.A. 1999. The PEDsQL: Measurement model for the pediatric quality of life inventory. Medical Care. 37(2):126-139.

World Health Organization 1948. Constitution of the World Health Organization. Basic Documents. WHO, Geneva

Developing a Draft Descriptive System

Developing the wording for the levels within the dimensions

6a.1 Introduction

The next two chapters describe how the dimensions identified in chapter 5 were developed into a descriptive system. Conventionally in instrument development, dimensions have multiple items and then levels are developed for these. Here, the dimensions only contain 1 item but need levels to be suitable for health state valuation work. There were two stages to this. This chapter describes how the wording for levels (response scales) within dimensions was developed and chapter 6b describes how the order of these levels was determined. This chapter begins by discussing approaches to level development taken by other instruments in the literature, looking at the issues of response scales, recall period, and respondent. It then reviews the scales used in existing paediatric generic measures before going on to describe the empirical work undertaken to develop the scales for this measure.

6a.2 Aim

The aim was to develop levels (response scales) for the dimensions identified from the interview work in order to construct a health state classification system.

6a.3 Background

6a.3.1 Dimension and item selection

Existing quality of life measures have generally taken an approach to descriptive system development whereby a series of items or statements are developed using focus groups, the literature or interviews. Work is then done to develop order and scales for these items, or response options could be based on Likert scale type responses (Streiner 1995). These are then reduced or sorted into factors or dimensions using psychometric techniques such as factor or Rasch analysis. Reduction of items is common as generally long lists of items are generated which are too long to have each item in the final questionnaire, hence testing is useful to identify redundant items (for example if items are not used or are very similar to another item), incomprehensible or ambiguous items and to test the internal consistency of a scale (Streiner 1995). Both factor analysis and Rasch techniques can be used, and used as complements rather than alternatives (Tennant 2004).

6a.3.2 Level Development

This work has taken a different approach, in that the interviews and analysis were driven by trying to determine dimensions of paediatric health related quality of life directly. The qualitative work provides supporting evidence as to why the dimensions arose and the terminology of the dimensions is based on the wording in the interviews. There is very little guidance in the literature about how to develop levels for dimensions directly. One way could be to consider the use of standard response scales.

6a.3.3 Scale options

Type

Some instruments use Likert type responses (Streiner 1995) which are based on options related to frequency (e.g. never, sometimes, often), the intensity/severity of a dimension (e.g. a little, moderately, a lot), or the level of agreement with something (strongly agree, disagree etc).

Existing generic preference based measures have taken different approaches when using scales. The EQ-5D takes the severity approach, using three levels for each dimension, the HUI2/3 has a mixture of both (severity and frequency) and the SF-36 (used to obtain the SF-6D) has a mixture of both, but is mainly a frequency based approach. The levels on the EQ-5D descriptive system, (a generic preference based measure for adults), were developed to be ordinal and were developed using an expert panel. The developers also recommend using severity based scales although they do not justify why (Kind 2005).

It can make a substantial difference to the descriptive system depending on the scale used. For example a frequency based scale may not capture the range of how something can affect a person, e.g. you can always be worrying, but only at a low level, which is different to being extremely worried. Equally, a scale based on severity may not adequately describe frequency. Another type of scale which is used in health status measures is level of agreement, which asks a respondent how much they agree (or disagree) with a statement. This type of scale does not really make sense for a preference based measure where a separate scale for each item level is not wanted. There is also a scale which asks you to indicate how much something bothers you however, again this is not suitable for a preference based measure. It is not useful for societal valuation, but may be useful for individual clinical decision making.

Response scales used in existing paediatric measures

One option in developing levels, for the dimensions developed in chapter 5 would be to use an existing paediatric scale. Existing generic paediatric measures were reviewed and Table 6a.1 below details the type of scales used (severity or frequency or agreement), together with the wording used in the scale.

Table 6a.1: Review of scales used in existing paediatric generic measures

QoL Measure	Type of scale	Example of wording used in the scale
Kidscreen	Severity Frequency	Not at all, slightly, moderately, very, extremely Never, seldom, quite often, very often, always
CHIP-CE	Frequency	No days, very few days, some days, almost every day, every day Never, almost never, sometimes, almost always, always

CHIP-AE	Frequency Level of agreement Other	Number of days a symptom or behaviour occurred in the past 4 weeks (5 point scale - no days to '15 to 28' days) Level of agreement, 4 point (completely agree to do not agree) Most recent occurrence, 5 point response. Number of occurrences in the past 12 months, 3 and 5 point response.
YQoL-R/S	Level of agreement	Not at all to completely 11 point Not at all to a great deal 11 point
TEDQL	How much you want to be/are like something	Really/a lot (like, good etc), a little bit, a little bit, really/a lot (dislike, etc)
Child quality of life questionnaire	Agreement Severity	7 point Likert scale As well as any other child of the same age up to the worst e.g. confined to bed Also some questions have a satisfaction response scale, and some a severity scale – extremely upset to not at all upset.
PEDS QoL	Frequency	3 point scale for 5-7 year olds not at all, sometimes, a lot 5 point Likert scale for 8 to 18 year olds never a problem, sometimes a problem, often a problem, always a problem

HUI 2/3 / HSCS – PS	Descriptors of severity	Descriptors HUI2: 3 to 5 levels of functioning HUI3: 5 or 6 levels of functioning
CHQ	Severity scale of limitations	4 point severity scale in how limited something is e.g. playing soccer or running (yes, limited a lot down to no, not limited).
Dartmouth COOP	Frequency Agreement	5 point Likert scales none of the time, a little of the time, some of the time, most of the time, all of the time yes, as much as I wanted, yes, quite a bit, yes, some, yes, a little, no, not at all
Exeter	Agreement	VAS anchored from very much like me to not very much like me Then again for don't want to be like that to really want to be like that
Functional Status II	Frequency	2 or 3 point Likert scales Never or rarely, some of the time, almost always Fully, partly, not at all
Generic Health Questionnaire	No details found	5 point Likert scale.
How are you	Frequency	4 point Likert scale example item in HTA review no, never, yes, sometimes, yes, often, yes, very often
Kindl	Frequency	5 point scale, never to always
Nordic QoL questionnaire	Above or below the base level	Not reported

Perceived Illness Experience	Agreement	5 point scales disagree to agree
Quality of Life Profile	Agreement Level of satisfaction	Importance and enjoyment ratings on a 5 point scale. Not at all important/no satisfaction at all to extremely important/extremely satisfied
TapQoL	Frequency	Some 0 to 2, some 0 to 4 scales, never, occasionally, often, occasionally or often, fine, not so good.
Warwick Child Health and Morbidity Profile	Mixture	4 categories of response
16D	Severity statements	Descriptive and includes health. 5 E.g. my state of health has no influence on my getting friends or being with friends My state of health makes getting friends or being with friends a little difficult The questions were formulated to measure the subject's function on each dimension however, when questions could be influenced by other non health factors, the questions were formulated to exclude the effect of non-health related factors on the function. (Apajasalo 1996)
17D	Severity statement then level of	Statement then 5 point response Not at all, a little, quite a lot, almost

	possibility agreement scale	impossible, totally impossible.
QWB	Descriptive statements	3 to 5 levels
TACQoL	Severity and frequency	3 and 4 point Likert scales yes, too little, never fine, not so good, quite bad, bad No, A bit, Very much

The majority of scales used are Likert type with a variety of response options and the vast majority are frequency based rather than severity. Most do not give any explanation on how the levels or scales were developed. Those with a shorter recall period, the 16D/17D and HUI2/3 are statement based (Apajasalo 1996) (Health Utilities Index).

6a3.4 Number of responses

There is little empirical work in the paediatric field with regard to the use of response options and children's ability to understand and use them across ages. Many existing measures use response options with between 3 and 7 points and there is literature which has shown that the minimum number of categories used by raters should be in the region of between 5 and 7 (Streiner 1995). Some measures use the same number of response options for each question, and some use different numbers of response options. The HUI2/3 and the 16D/17D use descriptive statements instead, however these are still ordinal (Apajasalo 1996) (Health Utilities Index). There are also developmental differences in children's ability to understand and respond to items on a Likert scale. Eight year old children can accurately use a 5 or 7 point scale to rate their health status whereas younger children tend to use more extreme responses. Some instruments have used visual aids to help with this, for example the Child Health and Illness Profile, which uses graduated circle sizes for the response options (Riley 2004).

ʒa.3.5 Recall period

The recall period is the time frame respondents are asked to think about when completing a questionnaire. In existing paediatric generic measures, there is a range of recall periods, from several weeks to the current day. More research is needed in this area about what is appropriate for children and different health conditions (Eiser 2001) (Matza 2004).

Many of the instruments based on a frequency approach ask questions about how often something has been the case over the past few weeks. The evidence from the qualitative interviews undertaken in this work indicates that children can recall information about their health and understand and describe it well, but often have difficulty remembering when they had a particular health problem or when an event had occurred. The advantage of asking about HRQoL today, is that you are focusing on a point in time and you also remove any potential problems with recall bias as children are thinking about the present time. The disadvantage is that this may miss important episodes in the context of a trial, particularly in episodic conditions.

6a.3.6 Format

Generally HRQoL measures take an approach of incorporating wording about health into the descriptive system to make sure it is based on health effects and not other factors, illustrated by Apajasalo M et al (1996) in their development of the 16D. In a HRQoL measure it can be argued that you want to capture information about problems related to health i.e. you do not just want to pick up for example that a child is worried because of an exam that day, so somewhere in the wording this needs to be incorporated. It could be incorporated in instructions at the beginning of the descriptive system, i.e. ask respondents to think about the questions in relation to their health today. Alternatively it could be part of each question, e.g. my health has affected my sleep a little.

It is also important to try and avoid negatively worded items in the format of the response options. These can be more confusing than positively worded items (Streiner 1995), and this is likely to be especially so for children. In particular, children may have difficult grasping the concept that they have to disagree with an item to indicate a positive answer, for example if the item said " I feel unwell much of the time", they would have to respond "No" to indicate a positive answer (Streiner 1995).

6a.3.7 Respondent

It could be argued that children are the population who know best how they feel and how they are and therefore ideally the respondent should be the child. Other generic paediatric measures are often proxy or interviewer administered. As the dimensions and wording are all based only on interviews with children, this should increase the likelihood of them being able to self complete. This is something that needs more research and is tested later on in the development of this measure. Work needs to be undertaken to determine whether they can do this or whether the measure needs to be interviewer or proxy administered and to what extent this depends on the age of the child.

6a3.8 Basis for deciding upon response scales

There were two options that could be taken to develop response scales:

1. Use one of the existing scales, however these are very varied and it is not obvious which is most appropriate.
2. Use the qualitative data obtained during the interview phase of this work to develop wording for the scales, looking at the phrases children use to describe the dimensions.

It was decided to return to the interview data in order to develop levels for the dimensions. Firstly to determine whether the dimensions are frequency based or severity based and secondly to develop the scales for the dimensions.

a.4 Methods

To determine whether the dimensions were to be based on frequency or severity, phrases were extracted from the interviews when the children were describing the dimensions and the manner in which something was described e.g. it's a bit annoying or it's quite annoying. This was used to guide whether each particular dimension was about severity or frequency. Phrases were extracted for all dimensions and for each age group separately.

Once it was established whether the dimensions were to be frequency or severity based, an approach based on the qualitative data was taken by developing a scale based on the wording used by the children in the interviews and using guidance on scale development from the methodological literature (Streiner 1995) together with what is required for a PBM (i.e. ordinal levels within each dimension) (Brazier 1999). The principles from the literature are as follows:

- Items should be clear, relevant and understandable
- Scales will be developed with 5-7 levels
- Language should be kept simple
- Double barrelled questions will be avoided (asking two different things within one question)
- Negatively worded items will be avoided, using positive wording styles instead
- Vague quantifiers will be avoided, although this can be very difficult in practice.

In addition, the following approach was also followed because of the use of qualitative data and the constraints of a PBM:

- The qualitative interviews will be used to guide the wording of the levels, by analysing how the children describe the problem, e.g. It hurts a bit, and it hurts a lot
- Levels will be ordinal, using an adjectival scale with discrete responses

- Language will be based on the qualitative data

6a.5 Results

Severity or frequency dimensions

The wording used to describe the dimensions are shown for each year group (Y3/Y4 and Y5/Y6) with the descriptive words underlined, in Appendix 6a.1. For every dimension, severity arose as the predominant characteristic. In a couple of dimensions (worrying and angry/annoyed/frustrated) frequency arose in one case in each. For *worrying*, this was a mixture of the two "I always get a bit worried". For *angry/annoyed/frustrated*, it was frequency "it's always annoying". For *sleep*, one child described it in frequency terms "can't get to sleep that often". In the *schoolwork, activities* and *daily routine* dimensions, children were describing how much they could or couldn't do something which indicates a severity approach.

As severity was the predominant characteristics for all dimensions, a severity based approach was taken to scale development.

Developing the wording for the dimension levels

The phrases used to describe the dimensions in the qualitative data are listed below by year group.

Severity wording from Y3/4 children

A bit, really, a little bit, **very, quite, kinda, a lot, quite a lot, much**,

Severity wording from Y5/6 children

A bit, kinda, quite, really, quite a lot, very, a lot, very much, **much**, at all

Those in bold text are common across age groups. As only one descriptive system was being produced for the year groups together (as noted in chapter

5), a combined wording pool was formed by putting all these terms together as follows:

Usable combined wording

at all	a little bit	a bit	quite	quite a lot
much	a lot	very	very much	really

The only phrase missed out is kinda, as this is a colloquial word.

Applying the wording to form a scale for each dimension

As noted in chapter 5, there are alternative terms used to describe the dimensions. Where more than one exists, alternative wordings were used to describe dimensions which were then tested as questions in the questionnaire piloting work that follows (chapters 7 and 8a). The dimensions and their alternative wording are shown in Table 5.8 (in chapter 5) and are reproduced here in Table 6a.2.

Table 6a.2: Dimensions of Health Related Quality of Life

	(Y3/Y4) (7-9 years)	(Y5/Y6) (9-11 years)
1	Worried Scared	Worried
2	Sad Upset	Sad Upset Unhappy Miserable
3	Annoyed Frustrated	Annoyed Frustrated Angry
4	Hurt Pain	Hurt Pain
5	School work	Learning
6	Daily Routine	Daily Routine

7	Tired Weak Drowsy	Tired Weak Energy Weary
8	Joining in activities that want to	Joining in activities that want to
9	Sleep	Sleep
10		Embarrassed
11	Jealous	

Not all terms were used as alternatives, as sometimes words were used by the older age group and so were more complex, for example miserable. As the questionnaire is being developed for the two age groups combined, where there was a choice over wording, the wording used by the younger age group was selected.

The final solution for testing was as follows. *Worried* and *scared* were developed as separate questions and *sad* and *upset* were developed as separate questions. *Miserable* is just a more sophisticated wording style by the older children and was therefore not included. *Unhappy* was felt not to be a good term for use in a questionnaire as it is negatively worded and so was not included. *Annoyed, frustrated* and *angry* were all developed as separate questions. *Hurt* and *pain* were developed as separate questions. *School work* and *learning* meant the same things in the interviews, therefore the younger children's terminology was used (i.e. *school work*). *Daily routine* was the same for both age groups so this was developed into a question. *Tired* and *weak* were developed into questions as *drowsy* and *weary* were not in common across age groups, and *energy* is the opposite meaning. *Joining in activities* was the same for both age groups so this was developed into a question. *Sleep* was the same for both age groups so this was developed into a question. Finally, *jealous* and *embarrassed* were both developed into questions.

This resulted in seventeen questions in total: *Worrying; Sad; Weak; Angry; Pain; Frustrated; Hurting; School Work; Upset; Tired; Annoyed; Scared; Sleep; Embarrassed; Jealous; Daily Routine* and *Joining in activities.*

The usable combined wording pool described previously was used to develop levels for each of the 17 questions. In addition, the wording used tried to incorporate the ways in which children had described the dimensions, for example for *worried, sad, angry, weak* and *embarrassed,* children were often using the term 'feel'. For *hurt* and *pain,* they were describing it in terms of it hurting or having pain.
This resulted in seven different types of scale, some of which were very similar, but had subtle differences depending on how the dimension fitted with the wording. The seven different scales and the questions to which they apply are listed below with an example of each scale given.

Scale 1 (Worrying, Sad, Weak, Angry, Frustrated, Upset, Tired, Annoyed, Scared, Embarrassed, Jealous)
I don't feel worried
I feel a little bit worried
I feel a bit worried
I feel quite worried
I feel very worried
I feel really worried

Scale 2 (Pain)
I don't have any pain
I have a little bit of pain
I have a bit of pain
I have quite a lot of pain
I have a lot of pain
I am really in pain

Scale 3 (Daily routine)

I have no problems with my daily routine

I have a few problems with my daily routine

I have some problems with my daily routine

I have many problems with my daily routine

I can't do my daily routine

Scale 4 (Hurting)

It doesn't hurt

It hurts a little bit

It hurts a bit

It hurts quite a bit

It hurts quite a lot

It hurts a lot

It really hurts

Scale 5 (Joining in activities)

I can join in with any of the activities that I want to

I can join in with most of the activities that I want to

I can join in with some of the activities that I want to I can join in with a few of the activities that I want to

I can join in with none of the activities that I want to

Scale 6 (Sleep)

My sleep is not affected

My sleep is a little bit affected

My sleep is a bit affected

My sleep is quite affected

My sleep is affected quite a lot

My sleep is really affected

My sleep is very affected

My sleep is affected a lot

I can't sleep at all

<u>Scale 7 (School Work)</u>

My school work is not affected
My school work is a little bit affected
My school work is a bit affected
My school work is quite affected
My school work is affected quite a lot
My school work is really affected
My school work is very affected
I can't do my school work

6a.6 Discussion and Conclusions

A draft descriptive system has been developed from the dimensions based on the qualitative data. The descriptive system is for both age groups combined. It contains 17 questions, some of which are alternative wording for the same dimensions, as the intention is to determine the best wording in subsequent testing. A scale could have been used from the paediatric literature however the only severity based scale in the literature for paediatric generic instruments is the scale from the KIDSCREEN (MAPI).

This scale is for children aged 8-18 years and uses the scale:

Not at all	slightly	moderately	very	extremely

The words slightly, moderately and extremely never appeared in the qualitative interviews undertaken in this research and seem difficult for young children to fully understand.

The dimensions contain levels which are based on severity empirically determined from the qualitative data. It could be argued that the majority of health conditions in the population interviewed were acute, therefore this is why the dimensions were described in severity terms, as you are more likely to be focussing on the current time rather than thinking about a condition you have always had and how it generally affects you. It is the case that there were more acute conditions in the interview sample, however there were many chronic conditions too such as hearing problems, vision problems, asthma, hay fever, allergies, weak wrists and ankles, eczema, hyperactive fits and abnormal muscle growth. Children with these problems described the dimensions mainly in terms of severity. The full list of health conditions by year group is shown at the end of chapter 5 in Appendix 5G and 5H.

The levels within each dimension are based on the wording used by children to describe the health problems, however the order of the levels, whilst perhaps appearing logical as set out above has not been verified by the children. Further work is needed to determine the order of the levels within the dimensions. The spacing of the scales is not necessarily even, however they do not have to be equally spaced as ultimately this will be a preference based instrument and those levels that are too close will drop out in future testing work. It is also likely that there are too many levels as whilst the principle was to aim for 5-7 levels, two of the scales have more than this number (sleep and school work with 9 and 8 respectively). In scale development however it is usual to start with too many levels and then reduce them. These issues are addressed in an empirical study reported in the next chapter.

Appendix 6a.1: Wording used in the interviews to describe the dimensions

Y3/Y4 Wording

1. Scared/worried

A bit scary

I felt nervous

you get really worried

A bit worried

I got really worried

I just get really worried

I felt like really, really worried

I was so scared

its so scary

makes me a bit scared

sometimes I worry

I get a bit scared sometime

I got a bit scared

It's scary

I was too scared

I get a bit scared

2. Sad/upset

I wasn't really happy

I feel sad

a bit sad

Unhappy

it's a bit sad

just sad

I felt a bit sad

I was a bit sad

Sad.

A little bit sad

it always makes me feel sad and upset

Upset and sad

Not very happy

Upset kind of

Very sad

very very upset

it made me a bit sad

I didn't feel that happy

not very happy

Sad

I felt erm quite sad

I wasn't really happy

Quite happy

Unhappy

3. Annoyed/frustrated

Annoying

It's just really frustrating

I just feel really frustrated

It was really annoying

bored and annoyed and irritated

I was really annoyed

Really really annoyed

it's quite annoying

really annoys me

they just kind of annoy me a bit.

it kinda annoys me

a bit annoyed.

It's a bit annoying.

it's annoying

it just is annoying

it still gets a bit annoying
it was really annoying
it's just annoying.

4. Hurting/pain

hurted a bit
it like really hurts
it hurts
it really hurts
really painful
really hurted
it just gets like pain
it just like hurts a lot
I get pains
it hurt so
it really hurted
it hurted quite a lot
it hurts much
It hurts a lot
really hurting
it really hurt.
it really hurts.
Painful.
sometimes it hurts a bit
it was painful,
It was really really painful.
they're painful.
it's painful
it like really hurts
It felt really really painful.
it hurt quite bad.
it's just a bit painful.

it hurt a bit.

5. Schoolwork/learning

It stopped me concentrating on my work

I can't read or write

It affected me working at school

it affected my writing

Harder to write

I can't write properly

It stopped me drawing

I couldn't write

I can't write properly

Cos I can't see and I don't know what to do.

I can't really listen properly

it stops me listening

it just like disturbs me

to be able to do work

I couldn't concentrate

I can't answer any answers at school or anything

6. Daily routine (eating, washing, dressing, getting ready)

It stopped me from drinking coca cola and that..

Well I could eat a little bit but then I didn't eat so much

The food won't go down and I didn't like it so, and I didn't want to eat.

I couldn't eat properly.

I couldn't eat properly it kept on going tight my nose when I ate, when I was swallowing I couldn't eat or anything

It stopped me drinking water

I have to sit down in the bath.

My mum's got to help me

I had to get up with my hands and try to sit down.

I couldn't drink.

I could only eat – drink stuff though a straw

I can't even eat

I can't clean my teeth properly

I had to crawl everywhere cos I couldn't stand up.

I just couldn't sit up

7. Feeling tired/don't feel like doing anything

I just feel really frustrated and tired and I can't do anything

I was so tired

I didn't' feel like getting up or anything

sometimes it makes me feel a bit tired.

I usually have to have like a rest

it just makes me really tired

I was really tired

you couldn't do anything at all

Just like tired and sleepy.

I was a bit weak,

I was a bit too weak

I felt really drowsy

I got quite a bit tired,

I feel quite tired

I didn't feel like getting up or anything

Sometimes I don't just want to do it.

I couldn't manage it.

I felt really drowsy

8. Able to join in activities that want to (e.g. playing out with friends, sports)

couldn't play with my friends

It stopped me playing,

it stopped me seeing my friends

Playing with my friends

It stops you going out and playing with your friends.

couldn't play with my friends

I don't get to play with my friends

I can join in with my friends.

I couldn't play with my friend.

I don't want to miss things and stuff.

I just couldn't do anything.

it stopped me from playing outside

Well I couldn't really play outside

It stops me running about and playing and stuff.

I can't even run and jump or anything

It stops me writing and running and skipping and

I had to miss out

it stopped me from playing outside

9. Sleep

but once or twice I woke up

waking up early

I didn't get much sleep

Sometimes I can't get to sleep really

I kept waking up

I can't get back to sleep

I couldn't go to sleep at night

I didn't have very much sleep

I couldn't get much sleep.

You can't get to sleep

keep me awake all night

I can't really get to sleep

it kind of affected me sleeping

I wouldn't get any sleep

it made me wake up early

10. Jealous

a bit jealous
you feel a bit jealous
I feel a bit jealous sometimes.
I wish mine weren't like that
I feel a bit jealous sometimes

Y5/Y6 Wording

1. Worrying

A bit worried
I always get a bit worried
I worry
I was a bit worried

2. Sad/unhappy

Feel upset
Feel sad
Felt unhappy
a bit sad
It makes me feel kinda sad.
It's making me feel quite sad
I just felt unhappy
I'm quite sad
I felt miserable
I felt upset
miserable
it makes me feel unhappy
I was sad
Quite sad
a bit upset

a bit miserable
Quite miserable

3. Angry/annoyed/frustrated

I get really frustrated.
It makes me feel angry
It's a bit annoying
it makes me a bit annoyed.
It was a bit annoying
which is quite annoying
it's quite annoying
It was annoying
it's always annoying
That's a bit annoying
it's just a bit annoying
It really annoyed me
Quite annoyed
it's just really annoying
I get really angry
really annoying
Get really frustrated
It's really annoying
Quite annoying
Felt frustrated

4. Hurting/itching/pain

Really sore
Dry sharp pain
Just hurts
Really hurt
Really painful
It hurts

It was sort of painful

very itchy

It just hurts

like a dry but sharp pain

it really hurts

really painfully

it really hurt

it hurts quite a lot

it hurt a lot

really hurting

It just hurts a bit

I was just really in pain

it hurt really badly

it would hurt a bit

They were really painful

really itchy

Really sore

really big horrible pain

5. Schoolwork/learning

Wouldn't be learning

Stopped me learning

I wouldn't be learning

you can't learn

affects learning

Missing the work

It stopped me learning.

I'm missing a bit of umm my education

6. Daily routine (eating, bathing, dressing, getting ready)

Moving round

Takes a bit longer

Couldn't really do very much

Couldn't do very much stuff

Stopped me eating

Couldn't eat as much

it takes a bit longer to get ready for school

7. Feeling weak/weary/tired/don't feel like doing anything

I don't feel like I have the energy to do anything.

Felt really weak

Didn't really have a lot of energy

Can't be bothered

didn't want to do anything

don't really feel like much

I feel like going to sleep

Really tired

I felt a bit weary

I felt really weak

I didn't really have a lot of energy

didn't have much energy

I just feel all weak

I was really tired

I don't feel like I have the energy to do anything.

I don't feel like doing anything.

I just can't be bothered with it

Well I just didn't want to do anything

I just was quite tired and didn't feel like doing much

you just wanna go to sleep

I don't want to do very much

I just felt like just laying in my bed and doing nothing

I feel like going to sleep, I'm tired.

Really tired.

I was really tired all the time

I was really tired.

I was very tired.

didn't feel like doing much

I just was quite tired

a bit tired

you just wanna go to sleep.

8. Able to join in activities that want to (e.g. playing out with friends, sports)

I have to stay inside and do nothing.

I have to stop and take a rest

I'm missing things that I normally do on the weekend

I wasn't able to join in with things that I normally do

I couldn't really do a lot of activities that took lots of energy

I couldn't do much

you can't do anything

you just have to sit out

we would have played.

being able to run around and stuff

I can't go

I couldn't go to the park

I can't do the things that I normally do

it stops me playing

I couldn't really play

I couldn't play out

I didn't get out and about as much

I can't run around or play football

I didn't get to play any sports.

I would find it harder than I find it now

I couldn't really do a lot of activities

I wouldn't go to anything

I wasn't able to join in with things that I normally do

I wasn't able to do that

not do all the activities they do but just like lie at home and do nothing

I couldn't do much

you can't really join in that much

I couldn't really do very much

it kinda stops you from going there

I can't do nothing

I can't go

I'm missing things that I normally do

Yeah I wasn't able to join in with things that I normally do

I didn't really do anything

you can't do some of the things that you want to do

you can't do anything

9. Sleep

I couldn't go to sleep.

couldn't go to sleep at all

I can't get to sleep properly

I couldn't sleep

I woke up once

I couldn't sleep at all

I wouldn't be able to get to sleep

I can't get to sleep properly

I would wake up

I can't go back to sleep

I woke up once

I don't really sleep that well cos I can't get to sleep

I can't sleep very well

I couldn't sleep at all

it woke me up in the night

I was awake all night

I couldn't get to sleep

I just stayed awake like all night.
I couldn't really get to sleep
I found it really hard to get to sleep
I found it quite hard to get to sleep
It's harder to get to sleep quicker
It affects my sleeping a bit
I can't get to sleep that often
I do get to sleep but it like takes a bit longer

10. Embarrassed/teased

I get a bit teased about that
to pick on me
people will laugh
you feel a bit embarrassed.
it was a bit embarrassing
I get a bit embarrassed
I'm embarrassed because some people laugh.
Embarrassed

References

Apajasalo, M., Sintonen, H., Holmberg, C., Sinkkonen, J., Aalberg, V., Pihko, H. et al. 1996. Quality of life in early adolescence: A sixteen-dimensional health-related measure (16D). Quality of Life Research. 5, 205-211.

Brazier, J.E., Deverill, M., Green, C., Harper, R. & Booth, A. 1999. A review of the use of health status measures in economic evaluation. 3[9]. Health Technology Assessment.

Eiser, C. & Morse, R. 2001. Quality-of-life measures in chronic diseases of childhood. [5;4]. Health Technology Assessment.

Health Utilities Index. http://healthutilities.biz/
Accessed 14/07/2008

Kind, P., Brooks, R. & Rabin, R. (Eds). 2005. EQ-5D concepts and methods, a developmental history. Chapter 3 pp32. Springer.

MAPI research trust. Patient-Reported Outcome and Quality of Life Instruments Database. Available from: www.proqolid.org Accessed 13/04/2006

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.

Riley, A. W., Forrest, C.B., Rebok, G.W., Starfield, B., Green, B.F., Robertson, J.A. et al. 2004. The Child Report Form of the CHIP-Child Edition: reliability and validity. Medical Care. 42(3), 221-231.

Streiner, D.L. & Norman, G.R. 1995. Health Measurement Scales. A Practical Guide to their development and use. Chapter 5. 2nd ed. Oxford University Press. Oxford.

Tennant. A., McKenna, S.P. & Hagell, P. 2004. Application of Rasch analysis in the development and application of quality of life instruments. Value in Health, 7 (Supplement 1), S22-S26.

Chapter 6b

Developing a Draft Descriptive System

Determining the order of the levels within each dimension

6b.1 Introduction

Whilst the scales developed in chapter 6a are based on children's descriptions, the ordinality of these scales has not been confirmed. As children have been involved at every stage of the development of this measure and the measure is intended for children, it is important to verify the order of the scales with them. This chapter describes how the order of the levels was determined, using rank methods with children.

6b.2 Aim

The aim was to determine the order of the levels within dimensions developed in the previous section, as judged by children.

6b.3 Methods

6b.3.1 Overview

This study tested out the ordinality of the scales developed in chapter 6a by asking children to rank statements in order of their severity. Children were sampled from the same two schools used in the original qualitative work. It was not necessary to obtain further consent as when consent was initially given, this was given for all stages of the fieldwork (qualitative interviews, ranking and testing of the draft descriptive system). As a year had elapsed since undertaking the qualitative work, all children had moved up a school year and there was a new intake of Y3 children. These children's parents were written to in exactly the same manner previously and asked to consent to their child being asked to take part in the ranking and/or the piloting of the questionnaire. Children undertook the ranking work in school and the results were used to determine the final order of the scales for the draft descriptive system.

b.3.2 Testing of scales

Levels were created for each wording of each dimension (henceforth termed a question) by applying the seven scales developed in chapter 6a. These scales were applied to seventeen questions in total: Worrying; Sad; Weak; Angry; Pain; Frustrated; Hurting; School Work; Upset; Tired; Annoyed; Scared; Sleep; Embarrassed; Jealous; Daily Routine; Joining in activities. Seventeen ranking exercises would be infeasible for each child to do in one sitting, so one question was selected for each of these seven scales for children to rank. This assumes that the ordinality of the scale is independent of the item (question).

The scales and their items tested in this study were:

I don't feel worried
I feel a little bit worried
I feel a bit worried
I feel quite worried
I feel very worried
I feel really worried

I don't have any pain
I have a little bit of pain
I have a bit of pain
I have quite a lot of pain
I have a lot of pain
I am really in pain

I have no problems with my daily routine
I have a few problems with my daily routine
I have some problems with my daily routine
I have many problems with my daily routine
I can't do my daily routine

It doesn't hurt
It hurts a little bit
It hurts a bit
It hurts quite a bit
It hurts quite a lot
It hurts a lot
It really hurts

I can join in with any of the activities that I want to
I can join in with most of the activities that I want to
I can join in with some of the activities that I want to I can join in with a few of the activities that I want to
I can join in with none of the activities that I want to

My sleep is not affected
My sleep is a little bit affected
My sleep is a bit affected
My sleep is quite affected
My sleep is affected quite a lot
My sleep is really affected
My sleep is very affected
My sleep is affected a lot
I can't sleep at all

My school work is not affected
My school work is a little bit affected
My school work is a bit affected
My school work is quite affected
My school work is affected quite a lot
My school work is really affected
My school work is very affected
I can't do my school work

Cards were created for each question being tested, with each card displaying a level and these were put together into a coloured envelope, generating seven envelopes, one for each dimension/scale. Children were given each envelope, one at a time, in a random order (chosen by them) and asked to rank the cards in order of severity (how bad they thought they were) from best to worst. Ties were allowed. Where children ranked cards as equal they were asked if they had a preference for the wording. The ranking work was first piloted on 10 children aged 7-11 years (5 male and 5 female). They were able to complete the tasks successfully and advised on the size of the cards, the font used and the colours of the card.

For the main study, 31 children were sampled from both schools involved in the research and each child carried out the seven ranking exercises. The aim of the sampling was to obtain an equal balance across gender and all year groups and to include both schools equally. Children were approached one by one and if they consented to the research, they were given the seven ranking tasks to do. All children carried the task out by themselves with the researcher sitting with them in the school library or the dining room. The children's rankings were recorded together with any comments on preferences for wording where levels were ranked equally.

6b.3.3 Analysis

The rank data was analysed by looking at the mean ranking and variation (standard deviation) and by using Kendall's coefficient of concordance test statistic. The approach of looking at the mean ranking is similar to work undertaken by Keller et al (1998) as part of their work testing the equivalence of translations of widely used response choice labels, where they looked at the mean response choice ratings by country and language.

The Kendall statistic is between 0 and 1 and is a measure of the agreement between rankings, 0 means there is no agreement between rankings. It measures the extent to which ordering by each of two (or more) variables would arrange the observations into the same numerical order (Bland 2001).

The rank data was coded using the mid rank method (Argyrous 2006) (Hinton 1995) as this is more appropriate for this type of analysis and ensures that the sum of ranks is maintained. That is, a rank of 1 was coded as 1, a rank of 2 was coded as 2 and where rankings were tied, each tied ranking was given a value of the midpoint of the previous and next ranks. For example, a ranking sequence where the second and third cards were ranked equally was coded as 1,2.5,2.5,4,5.

Where there was a very small difference between mean rankings, this was taken to indicate that only one statement was needed for the descriptive system. A mean ranking of less than 0.20 (chosen as a very low and conservative estimate) was taken to be a small difference. In order to choose between the statements, the variation and the preferences of children for the wording was examined, with the least amount of variation taking priority.

6b.4 Results

All 31 children consented to take part and all children completed all seven ranking tasks. The characteristics of the sample are shown in Table 6b.1:

Table 6b.1: Characteristics of the sample

Characteristic	N
Hunter's Bar Junior School	16
Firs Hill Community Primary School	15
Male	15
Female	16
Y3	8
Y4	8
Y5	8

Y6	7
White	17
Mixed/dual heritage	2
Asian or Asian British	12
Black or Black British	0
Chinese	0
Excellent health	10
Very Good health	11
Good health	9
Fair health	1
Poor health	0

Table 6b.2 shows the mean rank, standard deviation, minimum and maximum for each of the seven sets of scales.

Table 6b.2: Mean rank, standard deviation, minimum and maximum position for each set of statements

	Set	Mean	SD	Min	Max
I can join in with any of the activities that I want to	1	1.10	0.30	1	2
I can join in with most of the activities that I want to	1	2.02	0.49	1	3.5
I can join in with some of the activities that I want to	1	3.08	0.43	2	4
I can join in with a few of the activities that I want to	1	3.81	0.46	2	4
I can join in with none of the activities that I want to	1	5.00	0.00	5	5
My sleep is not affected	2	1.00	0.00	1	1
My sleep is a little bit affected	2	2.52	0.71	2	4
My sleep is a bit affected	2	2.77	0.59	2	4
My sleep is quite affected	2	3.82	0.75	2	6
My sleep is affected quite a lot	2	5.08	0.50	4	7
My sleep is very affected	2	7.23	0.92	5	8
My sleep is really affected	2	7.27	0.69	6	8
My sleep is affected a lot	2	6.31	0.76	4.5	8

I can't sleep at all	2	9.00	0.00	9	9
My school work is not affected	3	1.19	1.08	1	7
My school work is a little bit affected	3	2.52	0.70	2	4
My school work is a bit affected	3	2.84	0.66	2	5
My school work is quite affected	3	3.85	0.83	2	7
My school work is affected quite a lot	3	5.02	0.70	2	7
My school work is very affected	3	6.29	1.08	1	7
My school work is really affected	3	6.29	0.69	4	7
I can't do my school work	3	8.00	0.00	8	8
I don't feel worried	4	1.00	0.00	1	1
I feel a little bit worried	4	2.27	0.48	2	4
I feel a bit worried	4	3.00	0.55	2	4
I feel quite worried	4	3.73	0.60	2	4
I feel very worried	4	5.42	0.45	5	6
I feel really worried	4	5.58	0.45	5	6
I don't have any pain	5	1.00	0.00	1	1
I have a little bit of pain	5	2.29	0.42	2	3
I have a bit of pain	5	2.71	0.42	2	3
I have quite a lot of pain	5	4.29	0.48	4	5.5
I have a lot of pain	5	5.08	0.59	4	6
I am really in pain	5	5.63	0.66	4	6
I have no problems with my daily routine	6	1.00	0.00	1	1
I have a few problems with my daily routine	6	2.27	0.40	2	3
I have some problems with my daily routine	6	2.73	0.40	2	3
I have many problems with my daily routine	6	4.03	0.18	4	5
I can't do my daily routine	6	4.97	0.18	4	5
It doesn't hurt	7	1.00	0.00	1	1
It hurts a little bit	7	2.34	0.57	2	4
It hurts a bit	7	2.89	0.59	2	4
It hurts quite a bit	7	3.77	0.48	2	4
It hurts quite a lot	7	5.29	0.51	5	7
It hurts a lot	7	5.95	0.57	5	7
It really hurts	7	6.76	0.56	5	7

Table 6b.3 shows the mean rank order with the difference in mean rank order between the levels. A difference of less than 0.20 is shown in bold type.

Table 6b.3

Level	Mean rank order	Difference
I can join in with any of the activities that I want to	1.10	0.92
I can join in with most of the activities that I want to	2.02	1.06
I can join in with some of the activities that I want to	3.08	0.73
I can join in with a few of the activities that I want to	3.81	1.19
I can join in with none of the activities that I want to	5.00	
My sleep is not affected	1.00	1.52
My sleep is a little bit affected	2.52	0.26
My sleep is a bit affected	2.77	1.05
My sleep is quite affected	3.82	1.26
My sleep is affected quite a lot	5.08	1.23
My sleep is affected a lot	6.31	0.92
My sleep is very affected	7.23	**0.05**
My sleep is really affected	7.27	1.73
I can't sleep at all	9.00	
My school work is not affected	1.19	1.32
My school work is a little bit affected	2.52	0.32
My school work is a bit affected	2.84	1.02
My school work is quite affected	3.85	1.16
My school work is affected quite a lot	5.02	1.27
My school work is very affected	6.29	**0.00**
My school work is really affected	6.29	1.71
I can't do my school work	8.00	
I don't feel worried	1.00	1.27
I feel a little bit worried	2.27	0.73
I feel a bit worried	3.00	0.73
I feel quite worried	3.73	1.69
I feel very worried	5.42	**0.16**
I feel really worried	5.58	
I don't have any pain	1.00	1.29
I have a little bit of pain	2.29	0.42

I have a bit of pain	2.71	1.58
I have quite a lot of pain	4.29	0.79
I have a lot of pain	5.08	0.55
I am really in pain	5.63	
I have no problems with my daily routine	1.00	1.27
I have a few problems with my daily routine	2.27	0.45
I have some problems with my daily routine	2.73	1.31
I have many problems with my daily routine	4.03	0.94
I can't do my daily routine	4.97	
It doesn't hurt	1.00	1.34
It hurts a little bit	2.34	0.55
It hurts a bit	2.89	0.89
It hurts quite a bit	3.77	1.52
It hurts quite a lot	5.29	0.66
It hurts a lot	5.95	0.81
It really hurts	6.76	

Table 6b.4 shows the Kendall coefficient for each set.

Table 6b.4

Set	Kendall Coefficient
1	0.925
2	0.939
3	0.880
4	0.918
5	0.914
6	0.954
7	0.933

The Kendall coefficient is very high for all sets. The lowest is for set 3 (school work). An agreement of 0.81 to 1.00 is classed to be almost perfect agreement for the Kappa statistic, which is another statistical measure of agreement (Landis 1977).

The difference in the mean rank order is low (less than 0.20) for the following statements (highlighted in bold in Table 6b.3).

Statement	Difference
My sleep is very affected	
My sleep is really affected	0.05
My school work is very affected	
My school work is really affected	0.0
I feel very worried	
I feel really worried	0.16

This indicated that only one statement was needed for the descriptive system. The preferences of children when these statements were ranked equally are shown in Table 6b.5.

Table 6b.5: Preferences of children for wording

	Statement	Children's Preference (n preferring each statement)
1	My sleep is very affected	1
	My sleep is really affected	1
2	My school work is very affected	3
	My school work is really affected	2
3	I feel very worried	3
	I feel really worried	2

(1) Sleep: 'really' had a lower standard deviation and a smaller range (shown in Table 6b.2). The preferences of the children were equal. Therefore 'My sleep is really affected' was chosen.

(2) School work: 'really' had a lower standard deviation and a smaller range (shown in Table 6b.2). 'Very' has one more vote. Therefore 'My school work is really affected' was chosen.

(3) Worried: 'really' and 'very' had the same standard deviation and range (shown in Table 6b.2). 'Very' was preferred by one vote. Therefore 'I feel very worried' was chosen.

6b.5 Discussion

The ranking exercise worked well with children and they were successfully able to complete the tasks with a 100% completion rate. The ordering of the statements resulting from the analysis made sense at face value and there was very good agreement in the rankings by children. These results can now be used to develop the draft descriptive system by applying these scales to all 17 dimensions in order to form a draft descriptive system. Whilst the sample size was quite low in this study, the high agreement in rankings gives confidence in the results produced.

6b.6 Conclusion

This ranking work with children has determined the order of the scales for each question and the draft descriptive system is now ready for piloting in a paediatric population.

References

Argyrous, G. 2006. Rank-order tests for two or more samples. In Statistics for Research with a guide to SPSS. 2nd edition. Sage Publications. London.

Bland, M. & Peacock, J. 2001. Methods based on rank order. In Statistical questions in evidence-based medicine. Oxford Medical Publications. Oxford University Press.

Hinton, P. R. 1995. Statistics Explained. A Guide for Social Science Students. Routledge.

Keller, S. D., Ware, J. E., Gandek, B., Aaronson, N. K., Alonso, J., Apolone, G. et al. 1998. Testing the Equivalence of Translations of Widely Used Response Choice Labels: Results from the IQOLA Project. Journal of Clinical Epidemiology. 51 (11), 933-944.

Landis, J.R. & Koch, G. G. 1977. The measurement of observer agreement for categorical data. Biometrics. 33: 159-174.